

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:34:52 ; Search time 55.3333 Seconds  
(without alignments)  
64.831 Million cell updates/sec

Title: US-10-008-377a-2

Perfect score: 52  
Sequence: 1 GVLLWEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 444336

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: Genesep23Sep04:\*  
2: Genesep1980s:\*  
3: Genesep2000s:\*  
4: Genesep2001s:\*  
5: Genesep2002s:\*  
6: Genesep2003as:\*  
7: Genesep2003bs:\*  
8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	6	AAO22984 Human p37
2	46	88.5	9	5	AAE31409 Human VGR
3	46	88.5	9	6	AAO22987 Human p37
4	46	88.5	9	7	ADK65216 Human VEG
5	46	88.5	9	7	ADK65225 Human VEG
6	46	88.5	9	7	ADK65275 Human VEG
7	46	88.5	9	7	ADK65285 Human VEG
8	46	88.5	9	7	ADK65266 Human VEG
9	46	88.5	9	7	ADK65235 Human VEG
10	46	88.5	9	8	ADM97287 Cytoxic
11	42	80.8	10	8	ADM97302 Cytoxic
12	32	61.5	10	4	AAO34960 Human com
13	31	59.6	9	2	AAW37279 Cytoxic
14	31	59.6	10	3	AAW37279 Cytoxic
15	28	53.8	9	2	AAW37279 Cytoxic
16	28	53.8	9	2	AAW37279 Cytoxic
17	28	53.8	9	2	AAW37279 Cytoxic
18	28	53.8	9	2	AAW37279 Cytoxic
19	28	53.8	9	2	AAW37279 Cytoxic
20	28	53.8	9	2	AAW37279 Cytoxic
21	28	53.8	9	2	AAW37279 Cytoxic
22	28	53.8	9	2	AAW37279 Cytoxic
23	28	53.8	9	2	AAW37279 Cytoxic
24	28	53.8	9	2	AAW37279 Cytoxic
25	28	53.8	9	2	AAW37279 Cytoxic

26	28	53.8	10	2	AAV46028
27	28	53.8	10	2	AAV46431
28	28	53.8	10	2	AAV46511
29	28	53.8	10	4	AAU02657
30	28	53.8	10	5	ABP30929
31	28	53.8	10	5	ABP30941
32	28	53.8	10	7	ADA08574
33	28	53.8	10	7	ADA08586
34	28	53.8	10	7	ADP08929
35	28	53.8	10	7	ADP08917
36	28	53.8	10	7	ADG46204
37	28	53.8	10	7	ADG46216
38	28	53.8	10	7	ADG46216
39	27	51.9	9	2	AAE61778
40	27	51.9	10	7	ADL17633
41	26	50.0	8	4	AAE78541
42	26	50.0	8	5	ADE03135
43	26	50.0	9	4	AAE88982
44	26	50.0	9	4	AAE88597
45	26	50.0	9	6	ABR04892

#### ALIGNMENTS

RESULT 1  
AAO22984 standard; peptide: 10 AA.  
ID AAO22984 standard; peptide: 10 AA.  
XX AAO22984;  
AC  
XX  
DT 17-SEP-2003 (first entry)  
XX  
DE Human p376-85 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.  
XX  
KW HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
KW cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;  
KW oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
KW t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
KW p376-85.  
XX  
XX Homo sapiens.  
XX  
XX WO2003042243-A2.  
XX  
XX 22-MAY-2003.  
XX  
XX 14-NOV-2002; 2002WO-EP012764.  
XX  
XX 15-NOV-2001; 2001US-00008377.  
XX  
XX (NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.  
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX Gambacorti-Passerini C, Passoni L;  
XX WPI; 2003-441791/41.  
XX  
XX New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
XX useful for preparing a composition for treating ALK-positive lymphoma,  
XX neuroblastoma or ALK-expressing neoplasia.  
XX  
XX Claim 1; Page 7; 33pp; English.  
XX  
XX The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma  
XX kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a  
XX cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase  
XX which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase  
XX fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell  
XX lines. More than 50% of ALCL cases possess a t(2;5) chromosomal  
XX translocation that leads to the expression of the NPM/ALK fusion protein  
XX which forms a potent oncogene when constitutively activated. Translocated  
XX ALK is a widely expressed tumour-associated antigen characteristic of ALK

-positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The peptides of the invention demonstrate cytostatic activity and induce an MHC (major histocompatibility complex) class I restricted cytotoxic lymphocyte response against tumour cells expressing the NPM/ALK fusion protein. Hence, the peptides may be utilised during the treatment, via immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas and ALK-expressing neoplasias. In addition, the peptides may be used during gene therapy. The current sequence is that of the human p376-85 ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention

XX  
SQ Sequence 10 AA;

Query Match 100.0%; Score 52; DB 6; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
Db 1 GVLLWEIFSL 10

RESULT 2  
AAE31409  
ID AAE31409 standard; peptide; 9 AA.  
XX  
AC AAE31409;

DT 24-FEB-2003 (first entry)

DE Human VGR3 peptide #3.

XX Human; T-lymphocyte; vaccine; viral infection; gene therapy; cancer.

OS Homo sapiens.

PN WO200272627-A2.

PD 19-SEP-2002.

PF 11-MAR-2002; 2002WO-EP002666.

XX 09-MAR-2001; 2001US-0274250P.

PR 14-MAY-2001; 2001US-0290353P.

PR 18-MAY-2001; 2001US-0291610P.

XX (CALL-) CALLISTOGEN AG.

XX Wrede P, Walden P, Eichler-Wertens M, Filter M;

DR WPI; 2002-759836/82.

XX Providing, identifying or optimizing peptides for inducing cytotoxic T-lymphocytes and for treating cancer, comprises selecting conserved PT regions in antigenic proteins and identifying CD8+ T-cell epitopes in the protein.

XX Disclosure; Page 15; 32pp; English.

XX The invention relates to a method for providing, identifying or/and CC optimizing peptides which induce cytotoxic T-lymphocytes and to the uses of the obtained peptides for vaccination. The method is useful for CC providing, identifying and/or optimizing peptides that are useful in CC manufacturing a pharmaceutical composition for the induction of cytotoxic T-lymphocytes, and for the prevention, treatment or diagnosis of cancer CC or viral infections. The invention is also used in gene therapy. The CC present sequence is human VGR3 peptide used to illustrate the method of the invention

XX  
SQ Sequence 9 AA;

Query Match 88.5%; Score 46; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLLWEIFSL 10  
Db 1 VLLWEIFSL 9

RESULT 3  
AAO22987  
ID AAO22987 standard; peptide; 9 AA.  
XX  
AC AAO22987;

DT 17-SEP-2003 (first entry)

DE Human p377-85 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.

XX HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
KW cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;  
KW oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
KW t(2; 5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
KW p377-85.

XX Homo sapiens.

PN WO2003042243-A2.

PD 22-MAY-2003.

PF 14-NOV-2002; 2002WO-EP012764.

PR 15-NOV-2001; 2001US-0006377.

XX (MAST-) ISTR NAZ STUDIO E CURA DEI TUMORI.

PA (SUDD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Gambacorti-Passerini C, Pasoni L;

DR WPI; 2003-441791/41.

XX New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide, useful for preparing a composition for treating ALK-positive lymphoma, neuroblastoma or ALK-expressing neoplasia.

XX Claim 1; Page 7; 33pp; English.

XX The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a CC cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase CC which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase CC fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell CC lines. More than 50% of ALCL cases possess a t(2;5) chromosomal CC translocation that leads to the expression of the NPM/ALK fusion protein CC which forms a potent oncogene when constitutively activated. Translocated CC ALK is a widely expressed tumour-associated antigen characteristic of ALK CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The CC peptides of the invention demonstrate cytostatic activity and induce an CC MHC (major histocompatibility complex) class I restricted cytotoxic CC lymphocyte response against tumour cells expressing the NPM/ALK fusion CC protein. Hence, the peptides may be utilised during the treatment, via CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas CC and ALK-expressing neoplasias. In addition, the peptides may be used CC during gene therapy. The current sequence is that of the human p377-85 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention

XX  
SQ Sequence 9 AA;

Query Match 88.5%; Score 46; DB 6; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4  
ADK65216  
ID ADK65216 standard; peptide; 9 AA.  
XX  
AC ADK65216;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Human VEGFR-1-derived peptide #2 for anti-angiogenesis treatment.  
XX  
KW antiangiogenic; cyostatic; antiinflammatory; immunosuppressive;  
KW ophthalmological; gynecological; antiarteriosclerotic; virucide;  
KW hepatocytic; dermatological; anti-HIV; antidiabetic; antipsoriatic;  
KW antirheumatic; antiarthritic; antithyroid; immunization; angiogenesis;  
KW vascular endothelial growth factor; VEGF; neuropilin;  
KW placental growth factor; tumor; neoplasias; metastases; inflammation;  
KW autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;  
KW edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;  
KW rheumatoid arthritis; thyroiditis; diabetic retinopathy;  
KW transplant rejection; macular degeneration; neovascular glaucoma;  
KW hemangioma; angiofibroma.  
XX  
OS Homo sapiens.  
XX  
PN WO2003086450-A1.  
XX  
PD 23-OCT-2003.  
XX  
PF 11-APR-2003; 2003WO-CU000004.  
XX  
PR 15-APR-2002; 2002CU-00000076.  
XX  
PA (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.  
XX  
PI Bequet Romero M, Acevedo Castro BE, Gavilondo Cowley JV;  
PI Fernandez Molina LE, Lopez Ochoa O, Silva Rodriguez RDLG;  
PI Musachio Lasa A, Galban Rodriguez E, Vazquez Blomquist DM;  
XX  
DR WPI; 2003-833615/77.  
XX  
PT Active immunization against angiogenic proteins, useful for treating e.g.  
PT tumors and inflammation, particularly contains vascular endothelial  
PT growth factor or its receptor.  
XX  
PS Disclosure; Page 19; 53pp; Spanish.  
XX  
CC The invention relates to an active immunization against angiogenic  
CC proteins comprising administration of a vaccination composition (A),  
CC optionally containing an adjuvant, that comprises polypeptides (I)  
CC directly associated with an increase in angiogenesis, their variants, or  
CC their encoding polynucleotides (II). Angiogenesis-associated polypeptides  
CC are: members of the vascular endothelial growth factor (VEGF) family,  
CC especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of  
CC VEGF-B, or VEGF-C or -D; a (co-)receptor of VEGF, particularly VEGFR-1,  
CC 2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental  
CC growth factor. (A) is used for treatment or prevention of tumors in  
CC mammals, particularly humans but also farm animals and pets, also many  
CC other conditions associated with excessive angiogenesis, specifically  
CC malignant or benign neoplasias (and their metastases), acute or chronic  
CC inflammation, autoimmunity and eye diseases (claimed). Among the diseases  
CC that may be treated are arthritis, endometriosis, arteriosclerosis,  
CC edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,  
CC psoriasis, rheumatoid arthritis, thyroiditis, diabetic retinopathy,  
CC transplant rejection, macular degeneration, neovascular glaucoma,  
CC hemangioma and angiofibroma. The method destroys cells that are the  
CC source of angiogenic proteins, rather than just neutralizing the activity  
CC of such proteins (as in passive immunization). This sequence represents  
CC an immunisation peptide of the invention derived from the VEGF receptor  
CC proteins.  
XX  
SQ Sequence 9 AA;

Query Match 88.5%; Score 46; DB 7; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VLLMEIFSL 10  
| | | | |  
Db 1 VLLMEIFSL 9  
XX  
DE Human VEGFR-2-derived peptide #1 for anti-angiogenesis treatment.  
XX  
KW antiangiogenic; cyostatic; antiinflammatory; immunosuppressive;  
KW ophthalmological; gynecological; antiarteriosclerotic; virucide;  
KW hepatocytic; dermatological; anti-HIV; antidiabetic; antipsoriatic;  
KW antirheumatic; antiarthritic; antithyroid; immunization; angiogenesis;  
KW vascular endothelial growth factor; VEGF; neuropilin;  
KW placental growth factor; tumor; neoplasias; metastases; inflammation;  
KW autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;  
KW edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;  
KW rheumatoid arthritis; thyroiditis; diabetic retinopathy;  
KW transplant rejection; macular degeneration; neovascular glaucoma;  
KW hemangioma; angiofibroma.  
XX  
OS Homo sapiens.  
XX  
PN WO2003086450-A1.  
XX  
PD 23-OCT-2003.  
XX  
PF 11-APR-2003; 2003WO-CU000004.  
XX  
PR 15-APR-2002; 2002CU-00000076.  
XX  
PA (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.  
XX  
PI Bequet Romero M, Acevedo Castro BE, Gavilondo Cowley JV;  
PI Fernandez Molina LE, Lopez Ochoa O, Silva Rodriguez RDLG;  
PI Musachio Lasa A, Galban Rodriguez E, Vazquez Blomquist DM;  
XX  
DR WPI; 2003-833615/77.  
XX  
PT Active immunization against angiogenic proteins, useful for treating e.g.  
PT tumors and inflammation, particularly contains vascular endothelial  
PT growth factor or its receptor.  
XX  
PS Disclosure; Page 19; 53pp; Spanish.  
XX  
CC The invention relates to an active immunization against angiogenic  
CC proteins comprising administration of a vaccination composition (A),  
CC optionally containing an adjuvant, that comprises polypeptides (I)  
CC directly associated with an increase in angiogenesis, their variants, or  
CC their encoding polynucleotides (II). Angiogenesis-associated polypeptides  
CC are: members of the vascular endothelial growth factor (VEGF) family,  
CC especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of  
CC VEGF-B, or VEGF-C or -D; a (co-)receptor of VEGF, particularly VEGFR-1,  
CC 2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental  
CC growth factor. (A) is used for treatment or prevention of tumors in  
CC mammals, particularly humans but also farm animals and pets, also many  
CC other conditions associated with excessive angiogenesis, specifically  
CC malignant or benign neoplasias (and their metastases), acute or chronic  
CC inflammation, autoimmunity and eye diseases (claimed). Among the diseases  
CC that may be treated are arthritis, endometriosis, arteriosclerosis,  
CC edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,  
CC psoriasis, rheumatoid arthritis, thyroiditis, diabetic retinopathy,  
CC

CC transplant rejection, macular degeneration, neovascular glaucoma,  
 CC hemangioma and angiofibroma. The method destroys cells that are the  
 CC source of angiogenic proteins, rather than just neutralizing the activity  
 CC of such proteins (as in passive immunization). This sequence represents  
 CC an immunisation peptides of the invention derived from the VEGF receptor  
 CC proteins.  
 XX  
 XX Sequence 9 AA;  
 SQ  
 Query Match 88.5%; Score 46; DB 7; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 VLWEIFSL 10  
 Db 1 VLWEIFSL 9  
 RESULT 6  
 ADK65275  
 ID ADK65275 standard; peptide; 9 AA.  
 XX  
 AC ADK65275;  
 DT 06-MAY-2004 (first entry)  
 DE Human VEGFR-2-derived peptide #11 for anti-angiogenesis treatment.  
 XX  
 XX antiangiogenic; cytostatic; antiinflammatory; immunosuppressive;  
 KW ophthalmological; gynecological; antiarteriosclerotic; vincicide;  
 KW hepatotropic; dermatological; anti-HIV; antidiabetic; antipsoriatic;  
 KW antirheumatic; antiarthritic; antithyroid; immunization; angiogenesis;  
 KW vascular endothelial growth factor; VEGF; neuropilin;  
 KW placental growth factor; tumor; neoplasias; metastases; inflammation;  
 KW autoimmunity; eye disease; diabetes; endometriosis; arteriosclerosis;  
 KW edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;  
 KW rheumatoid arthritis; thyroiditis; diabetic retinopathy;  
 KW transplant rejection; macular degeneration; neovascular glaucoma;  
 KW hemangioma; angiofibroma.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO2003086450-A1.  
 PN  
 XX 23-OCT-2003.  
 PD  
 XX 11-APR-2003; 2003WO-CU000004.  
 PF  
 XX 15-APR-2002; 2002CU-00000076.  
 PR  
 XX (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.  
 PA  
 XX Bequet Romero M, Acevedo Castro BE, Gavilondo Cowley JV;  
 PI Fernandez Molina LE, Lopez Oceljo O, Silva Rodriguez RDC;  
 PI Musachio Lasa A, Galban Rodriguez E, Vazquez Blomquist DM;  
 DR MPI, 2003-833615/77.  
 XX  
 XX Active immunization against angiogenic proteins, useful for treating e.g.  
 PT tumors and inflammation, particularly contains vascular endothelial  
 PT growth factor or its receptor.  
 PT  
 XX Disclosure; Page 19; 53pp; Spanish.  
 XX  
 XX The invention relates to an active immunization against angiogenic  
 CC proteins comprising administration of a vaccination composition (A),  
 CC optionally containing an adjuvant, that comprises polypeptides (I),  
 CC directly associated with an increase in angiogenesis, their variants, or  
 CC their encoding polynucleotides (II). Angiogenesis-associated polypeptides  
 CC are: members of the vascular endothelial growth factor (VEGF) family,  
 CC especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of  
 CC VEGF-B, or VEGF-C or -D; a (co-)receptor of VEGF, particularly VEGFR-1,  
 CC 2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental

CC growth factor. (A) is used for treatment or prevention of tumors in  
 CC mammals, particularly humans but also farm animals and pets, also many  
 CC other conditions associated with excessive angiogenesis, specifically  
 CC malignant or benign neoplasias (and their metastases), acute or chronic  
 CC inflammation, autoimmunity and eye diseases (claimed). Among the diseases  
 CC that may be treated are arthritis, endometriosis, arteriosclerosis,  
 CC edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,  
 CC psoriasis, rheumatoid arthritis, thyroiditis, diabetic retinopathy,  
 CC transplant rejection, macular degeneration, neovascular glaucoma,  
 CC hemangioma and angiofibroma. The method destroys cells that are the  
 CC source of angiogenic proteins, rather than just neutralizing the activity  
 CC of such proteins (as in passive immunization). This sequence represents  
 CC an immunisation peptides of the invention derived from the VEGF receptor  
 CC proteins.  
 XX  
 XX Sequence 9 AA;  
 SQ  
 Query Match 88.5%; Score 46; DB 7; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 VLWEIFSL 10  
 Db 1 VLWEIFSL 9  
 RESULT 7  
 ADK65285  
 ID ADK65285 standard; peptide; 9 AA.  
 XX  
 AC ADK65285;  
 DT 06-MAY-2004 (first entry)  
 DE Human VEGFR-3-derived peptide #11 for anti-angiogenesis treatment.  
 XX  
 XX antiangiogenic; cytostatic; antiinflammatory; immunosuppressive;  
 KW ophthalmological; gynecological; antiarteriosclerotic; vincicide;  
 KW hepatotropic; dermatological; anti-HIV; antidiabetic; antipsoriatic;  
 KW antirheumatic; antiarthritic; antithyroid; immunization; angiogenesis;  
 KW vascular endothelial growth factor; VEGF; neuropilin;  
 KW placental growth factor; tumor; neoplasias; metastases; inflammation;  
 KW autoimmunity; eye disease; diabetes; endometriosis; arteriosclerosis;  
 KW edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;  
 KW rheumatoid arthritis; thyroiditis; diabetic retinopathy;  
 KW transplant rejection; macular degeneration; neovascular glaucoma;  
 KW hemangioma; angiofibroma.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO2003086450-A1.  
 PN  
 XX 23-OCT-2003.  
 PD  
 XX 11-APR-2003; 2003WO-CU000004.  
 PF  
 XX 15-APR-2002; 2002CU-00000076.  
 PR  
 XX (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.  
 PA  
 XX Bequet Romero M, Acevedo Castro BE, Gavilondo Cowley JV;  
 PI Fernandez Molina LE, Lopez Oceljo O, Silva Rodriguez RDC;  
 PI Musachio Lasa A, Galban Rodriguez E, Vazquez Blomquist DM;  
 DR MPI, 2003-833615/77.  
 XX  
 XX Active immunization against angiogenic proteins, useful for treating e.g.  
 PT tumors and inflammation, particularly contains vascular endothelial  
 PT growth factor or its receptor.  
 PT  
 XX Disclosure; Page 19; 53pp; Spanish.  
 XX  
 XX The invention relates to an active immunization against angiogenic



CC	protein comprising administration of a vaccination composition (A),
CC	optionally containing an adjuvant, that comprises polypeptides (I)
CC	directly associated with an increase in angiogenesis; their variants, or
CC	their encoding polynucleotides (II). Angiogenesis-associated polypeptides
CC	are: members of the vascular endothelial growth factor (VEGF) family,
CC	especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of
CC	VEGF-B, or VEGF-C or -D; a (co-)receptor of VEGF, particularly VEGFR-1, -
CC	2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental
CC	growth factor. (A) is used for treatment or prevention of tumors in
CC	mammals, particularly humans but also farm animals and pets, also many
CC	other conditions associated with excessive angiogenesis, specifically:
CC	malignant or benign neoplasias (and their metastases), acute or chronic
CC	inflammation, autoimmunity and eye diseases (claimed). Among the diseases
CC	that may be treated are arthritis, endometriosis, arteriosclerosis,
CC	psoriasis, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,
CC	osteoarthritis, rheumatoid arthritis, thyroiditis, diabetic retinopathy,
CC	transplant rejection, macular degeneration, neovascular glaucoma,
CC	hemangioma and angiodysplasia. The method destroys cells that are the
CC	source of angiogenic proteins, rather than just neutralizing the activity
CC	of such proteins (as in passive immunization). This sequence represents
CC	an immunisation peptides of the invention derived from the VEGF receptor
CC	proteins.
CC	
XX	
SQ	Sequence 9 A/;
CY	Query Match 88.5%; Score 46; DB 7; Length 9;
Dn	Best Local Similarity 100.0%; Pred.No. 1.7e+06;
	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0
CY	2 VLMEIFSL 10         1 VLMEIFSL 9
Db	1 VLMEIFSL 9
RESULT 8	
ID	ADK65266 standard; peptide; 9 AA.
XX	
AC	ADK65266;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Human VEGFR-1-derived peptide #12 for anti-angiogenesis treatment.
XX	
KW	antiangiogenic; cytostatic; antiinflammatory; immunosuppressive;
KW	ophthalmological; gynecological; antiatherosclerotic; vincide;
KW	hepatotropic; dermatological; anti-HIV; antididiabetic; antipsoriatic;
KW	antineuraptic; antiarthritic; antithyroid; immunization; angiogenesis;
KW	vascular endothelial growth factor; VEGF; neuropilin;
KW	placental growth factor; tumor; neoplasias; metastases; inflammation;
KW	autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;
KW	edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;
KW	rheumatoid arthritis; thyroiditis; diabetic retinopathy;
KW	transplant rejection; macular degeneration; neovascular glaucoma;
KW	hemangioma; angiodysplasia.
OS	Homo sapiens.
XX	
WO	WO2003086450-A1.
PD	
FD	23-OCT-2003.
XX	
FF	11-APR-2003; 2003WO-CU000004.
XX	
PR	15-APR-2002; 2002CU-0000076.
XX	
PA	(INGG-) CENT ING GENETICA & BIOTECNOLOGIA.
PI	Bernier Romero M., Acevedo Castro BE, Gallardo Cowley JV;
PI	Fernandez Molina LE, Lopez Ojejo O, Silva Rodriguez RLC,
PI	Musacchio Laeta A, Galban Rodriguez E, Vazquez Blomquist DW,
OR	WI.; 2003-833615/77.

XX Active immunization against angiogenic proteins, useful for treating e.g.  
PT tumors and inflammation, particularly contains vascular endothelial  
PT growth factor or its receptor.  
PS Disclosure, Page 19, 53pp; Spanish.  
XX The invention relates to an active immunization against angiogenic  
CC proteins comprising administration of a vaccination composition (A),  
CC optionally containing an adjuvant, that comprises polypeptide (I),  
CC directly associated with an increase in angiogenesis, their variants, or  
CC their encoding polynucleotides (II). Angiogenesis-associated polypeptides  
CC are: members of the vascular endothelial growth factor (VEGF) family,  
CC especially the 121, 165 or 183 isoforms of VEGF-A, the 167 isoform of  
CC VEGF-B, or VEGF-C or -D; a (co-)receptor of VEGF, particularly VEGFR-1, -  
CC 2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental  
CC growth factor. (A) is used for treatment or prevention of tumors in  
CC mammals, particularly humans but also farm animals and pets, also many  
CC other conditions associated with excessive angiogenesis, specifically  
CC malignant or benign neoplasias (and their metastases), acute or chronic  
CC inflammation, autoimmunity and eye diseases (claimed). Among the diseases  
CC that may be treated are arthritis, endometriosis, arteriosclerosis,  
CC edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,  
CC psoriasis, rheumatoid arthritis, thyroiditis, diabetic glaucoma,  
CC transplant rejection, macular degeneration, neovascular retinopathy,  
CC hemangioma and angiodioma. The method destroys cells that are the  
CC source of angiogenic proteins, rather than just neutralizing the activity  
CC of such proteins (as in passive immunization). This sequence represents  
CC an immunisation peptides of the invention derived from the VEGF receptor  
CC proteins.  
XX  
SQ Sequence 9 AA:  
Query Match 88.5%; Score 46; DB 7; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1,7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VLIMEIFSL 10  
|||  
Db 1 VLIMEIFSL 9  
RESULT 9  
ADK65235  
ID ADK65235 standard; peptide; 9 AA.  
AC  
XX ADK65235;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Human VEGFR-3-derived peptide #1 for anti-angiogenesis treatment.  
XX  
KW antiangiogenic; cytostatic; antiinflammatory; immunosuppressive;  
KW ophthalmological; gynecological; antiarteriosclerotic; vitucide;  
KW hepatocytic; dermatological; anti-HIV; antidiabetic; antipsoriatic;  
KW antithematic; antiarthritic; antithyroid; immunization; angiogenesis;  
KW vascular endothelial growth factor; VEGF; neuropilin;  
KW placental growth factor; tumor; neoplasias; metastases; inflammation;  
KW autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;  
KW edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;  
KW rheumatoid arthritis; thyroiditis; diabetic retinopathy;  
KW transplant rejection; macular degeneration; neovascular glaucoma;  
KW hemangioma; angiodioma.  
XX  
OS Homo sapiens.  
XX  
PN MO2003086450-A1.  
XX  
PD 23-OCT-2003.  
XX  
FF 11-APR-2003; 2003WO-CU000004.  
XX  
FR 15-APR-2002; 2002CU-00000076.

XX (ING-) CENT ING GENETICA & BIOTECNOLOGIA.

XX Benget Romero M, Acevedo Castro BE, Gavilondo Cowley JV;

XX Fernandez Molina LE, Lopez Ocojo O, Silva Rodriguez PDC;

XX Musachio Lasa A, Galban Rodriguez E, Vazquez Blomquist DM;

XX WPI; 2003-833615/77.

XX Active immunization against angiogenic proteins, useful for treating e.g.

XX tumors and inflammation, particularly contains vascular endothelial

XX growth factor or its receptor.

XX Disclosure; Page 19; 53pp; Spanish.

XX The invention relates to an active immunization against angiogenic

XX proteins comprising administration of a vaccination composition (A),

XX optionally containing an adjuvant, that comprises polypeptides (1)

XX directly associated with an increase in angiogenesis, their variants, or

XX their encoding polynucleotides (II). Angiogenesis-associated polypeptides

XX are: members of the vascular endothelial growth factor (VEGF) family,

XX especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of

XX VEGF-B, or VEGF-C or -D, a (co-)receptor of VEGF, particularly VEGFR-1,

XX 2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental

XX growth factor. (A) is used for treatment or prevention of tumors in

XX mammals, particularly humans but also farm animals and pets, also many

XX other conditions associated with excessive angiogenesis, specifically

XX malignant or benign neoplasias (and their metastases), acute or chronic

XX inflammation, autoimmunity and eye diseases (claimed). Among the diseases

XX that may be treated are arthritis, endometriosis, arteriosclerosis,

XX edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,

XX psoriasis, rheumatoid arthritis, thyroiditis, diabetic retinopathy,

XX transplant rejection, macular degeneration, neovascular glaucoma,

XX hemangioma and angiodysplasia. The method destroys cells that are the

XX source of angiogenic proteins, rather than just neutralizing the activity

XX of such proteins (as in passive immunization). This sequence represents

XX an immunisation peptides of the invention derived from the VEGF receptor

XX proteins.

XX Sequence 9 AA;

XX

XX

XX

XX

XX

XX

PR 07-MAR-2003; 2003JP-00062003.

PR 11-JUN-2003; 2003JP-00167042.

XX (ONCO-) ONCOTHERAPY SCI INC.

XX (TAHA/) TAHARA H.

XX Tahara H, Wada S, Tsunoda T;

XX WPI; 2004-295074/27.

XX Nonapeptides comprising cytotoxic T-cell inducing activity, useful as

XX vaccine for treating proliferation of malignant tumors.

XX Example 6; Page 82; 100pp; Japanese.

XX The present invention relates to nona- and decapeptides comprising

XX cytotoxic T-cell (CTL) inducing activity which are KDR peptides binding

XX to HLA-A2402, containing partial sequences of receptor protein VEGFR-2

XX (Vascular endothelial growth factor receptor 2), or derived from these

XX sequences by addition or substitution of one or more amino acid residues.

XX The peptides can be used to produce a cancer vaccine for treatment and

XX prevention of cancer (including stomach, duodenum, colon, lung, breast,

XX prostate and brain cancer) and suppression of proliferation and

XX metastasis of malignant tumors. The vaccine is also useful for treatment

XX of diabetic retinopathy, chronic rheumatoid arthritis and

XX atherosclerosis. The present sequence is a peptide of the invention.

XX Sequence 9 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Example 6; Page 87; 100pp; Japanese.

CC The present invention relates to nona- and decapeptides comprising  
CC cytotoxic T-cell (CTL) inducing activity which are KDR peptides binding  
CC to HLA-A2402, containing partial sequences of receptor protein VEGFR-2  
CC (Vascular endothelial growth factor receptor 2), or derived from these  
CC sequences by addition or substitution of one or more amino acid residues.  
CC The peptides can be used to produce a cancer vaccine for treatment and  
CC prevention of cancer (including stomach, duodenum, colon, lung, breast,  
CC prostate and brain cancer) and suppression of proliferation and  
CC metastasis of malignant tumors. The vaccine is also useful for treatment  
CC of diabetic retinopathy, chronic rheumatoid arthritis and  
CC atherosclerosis. The present sequence is a peptide of the invention.

XX Sequence 10 AA;

Query Match 80.8%; Score 42; DB 8; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.98;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LLMWFSL 10  
|||  
1 LLMWFSL 8

Db 1 LLMWFSL 8

RESULT 12

AAAG94960  
ID AAG94960 standard; peptide; 10 AA.

AC AAG94960;

DT 18-SEP-2001 (first entry)

DE Human complementary peptide, SEQ ID NO: 1154.

KW Human; complementary peptide; ligand; drug discovery; drug design.

OS Homo sapiens.

FN WO200142277-A2.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB004776.

PR 13-DEC-1999; 99GB-00029464.

PA (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

DR WPI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides to  
XX proteins encoded by genes of the human genome; useful in an assay for  
XX screening and identifying of one or more novel peptides which are drug  
XX candidates or pro-drugs.

XX Example 4; Page 205; 646pp; English.

CC The invention relates to a set of complementary peptide ligands generated  
CC from the human genome. The complementary peptides interact with their  
CC relevant target proteins encoded in the human genome. They can be used as  
CC reagents in drug discovery and as lead ligands to facilitate drug design  
CC and development. The present sequence is a complementary peptide provided  
CC in the specification

XX Sequence 10 AA;

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 83.3%; Pred. No. 56;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 5 WEIIFSL 10  
|||  
1 WEIIFSL 6

Db 1 WEIIFSL 6

RESULT 13

AAW97279 standard; peptide; 9 AA.

AC AAW97279;

DT 17-OCT-2003 (revised)

DT 27-AUG-2003 (revised)

DT 20-MAR-2003 (revised)

DT 10-MAY-1999 (first entry)

DE Cytotoxic T-cell epitope from antigen gp85 of EBV.

KW Cytotoxic T-cell epitope; EBV; structural antigen; EBV infection;  
KW infectious mononucleosis; IM; NPC; HD; Burkitt's lymphoma;  
KW post transplantation lymphoproliferative disease; vaccine.

OS Human herpesvirus 4.

FN WO9902550-A1.

PD 21-JAN-1999.

PF 10-JUL-1998; 98WO-AU000531.

PR 10-JUL-1997; 97AU-00007841.

PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

PA (HALT-) HALT INST MEDICAL RES WALTER & ELIZA.

PA (CSLC-) CSL LTD.

PA (SHER/) SHERITT M A.

PI Burrows SR, Khanna R;

DR WPI; 1999-120774/10.

PF New Epstein-Barr virus antigens - comprise cytotoxic T-cell epitopes,  
PF used to treat or protect against EBV infection, e.g. in nasopharyngeal  
PF carcinoma or Hodgkin's disease.

XX Example 2; Page 28; 73pp; English.

CC AAW97258-90 represent cytotoxic Epstein-Barr virus (EBV) T-cell epitopes  
CC derived from an EBV structural antigen. The epitopes (and nucleic acids  
CC encoding them) can be used for reducing the risk of EBV infection in a  
CC subject. They can be used for reducing the risk of infectious  
CC mononucleosis (IM) or post transplantation lymphoproliferative disease in  
CC a subject. The epitopes can also be used for treating or preventing NPC  
CC or HD or Burkitt's lymphoma. They can also be used in subunit vaccines  
CC with other antigens, e.g. tetanus toxoid, diphtheria toxoid, Bordetella  
CC pertussis antigens, poliovirus antigens, purified protein derivative  
CC (PPD), gp350 protein, helper epitopes or combinations. (Updated on 20-MAR  
CC -2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)  
CC (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 9 AA;

Query Match 59.6%; Score 31; DB 2; Length 9;  
Best Local Similarity 83.3%; Pred. No. 1.7e+06;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 VLLMEI 7  
|||  
4 VLLMEV 9

Db 4 VLLMEV 9

```

RESULT 14
AAB19600 standard; peptide; 10 AA.
XX
XX AAB19600;
XX
XX 22-UTAN-2001 (first entry)
XX
XX Human CASB616 epitope binding HLA_A0201.
XX
XX CASB616; EPHB2; ERK; EPH3; EPH3; DRT; HEK5; EPHB2V;
XX receptor protein tyrosine kinase; human; epitope; colon cancer;
XX ovary cancer; tumour; autoimmune disease; vaccine; therapy; diagnosis;
XX HLA-A0201.
XX
XX Homo sapiens.
XX
XX WO200053216-A2.
XX
XX 14-SEP-2000.
XX
XX 26-FEB-2000; 2000WO-EP001587.
XX
XX 05-MAR-1999; 99GB-00005124.
XX
XX (SMIR ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Vinals Y De Bassols;
XX
XX WPI; 2000-587384/55.
XX
XX Vaccine composition for treating ovarian and colon cancer, comprises
XX CASB616 polypeptides, polynucleotides or antigen presenting cells
XX expressing the polypeptides.
XX
XX Example 6; Page 39; 57pp; English.
XX
XX The present sequence is that of a human CASB616 epitope that binds the
XX HLA-A0201 allele. The epitope corresponds to amino acid residues 810-819
XX of the CASB616 polypeptide (see AAB19590 and AAB19591). It is 1 of 10
XX epitopes (see AAB19592) identified during an analysis of human cellular
XX immune responses to CASB616. CASB616 is a tumour specific antigen.
XX CASB616 polypeptides and polynucleotides are used in claimed vaccine
XX compositions for specific prophylactic or therapeutic immunization
XX against tumours, especially colon cancer (claimed) and ovarian cancer, or
XX against autoimmune disease. They can also be used to diagnose the
XX occurrence of tumour cells
XX
XX Sequence 10 AA;
SQ
Query Match 59.6%; Score 31; DB 3; Length 10;
Best Local Similarity 42.9%; Pred. No. 84;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVTLMWEI 7
DB 4 GVTLMWEI 10

```

```

RESULT 15
AAY46428
ID AAY46428 standard; peptide; 9 AA.
XX
XX AAY46428;
XX
XX 01-DEC-1999 (first entry)
XX
XX Immunogenic peptide having a human leukocyte antigen binding motif #1039.
XX
XX Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
XX immune response; T cell activation; major histocompatibility complex;
XX cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
XX

```

```

XX prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
XX vaccine; immunisation.
XX
XX Synthetic.
XX
XX Homo sapiens.
XX
XX WO9945954-A1.
XX
XX 16-SEP-1999.
XX
XX 13-MAR-1998; 98WO-US005039.
XX
XX 13-MAR-1998; 98WO-US005039.
XX
XX (EPIM-) EPIMUNE INC.
XX
XX Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX WPI; 1999-551214/46.
XX
XX New immunogenic peptides with HLA binding motif, useful in treatment and
XX diagnosis of cancers and viral diseases.
XX
XX Claim 1; Page 71; 150pp; English.
XX
XX AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
XX having a human major histocompatibility complex (MHC) Class I (also known
XX as human leukocyte antigen (HLA)) binding motif. The immunogenic peptides
XX can bind to a specific HLA allele (i.e. HLA-A subtypes HLA-A2.1, A1, A3.2
XX or A24.1 or HLA-B or C) and induce a cytotoxic T cell response against
XX the antigen from which the peptide is derived. Cytotoxic T lymphocytes
XX (CTLs) which destroy antigen-bearing cells are normally induced by an
XX antigen in the form of a peptide fragment bound to a HLA molecule, rather
XX than the intact foreign antigen itself, and are particularly important in
XX tumour rejection and in fighting viral infections. The peptides are
XX therefore useful therapeutically to treat or prevent viral infections and
XX cancers in mammals (especially humans) e.g. prostate cancer, hepatitis B
XX and C, AIDS, and renal carcinoma. They can be administered as vaccines to
XX elicit an immune response in individuals susceptible or otherwise at risk
XX of viral infection or cancer, or used to treat chronic or acute
XX conditions. They are also useful diagnostically, and can be used to
XX induce a cytotoxic T cell response, by contacting a cytotoxic T cell with
XX the peptide e.g. to produce CTLs ex vivo for infusion back into a
XX patient. The polynucleotides encoding the immunogenic peptides are also
XX useful therapeutically and for immunisation as above
XX
XX Sequence 9 AA;
SQ
Query Match 53.8%; Score 28; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.7e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 GVTLMWEI 7
DB 3 GVTLMWEI 9

```

Search completed: November 30, 2004, 08:09:35  
Job time : 58.333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 20 Seconds  
(without alignments)

48.108 Million cell updates/sec

Title: US-10-008-377A-1

Perfect score: 47

Sequence: 1 SLAMDILHV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description		
1	47	100.0	680	4	138491	nucleophosmin/anap
2	47	100.0	1621	4	T30200	protein-tyrosine k
3	37	78.7	311	2	S65091	geranylgeranyltran
4	36	76.6	1039	2	F71427	hypothetical prote
5	36	76.6	4717	2	T41581	hypothetical coila
6	35	74.5	232	2	H69294	conserved hypotet
7	35	74.5	441	2	A41591	endothelin recepto
8	35	74.5	441	2	S13425	endothelin recepto
9	35	74.5	442	1	U71042	endothelin recepto
10	35	74.5	442	2	I57950	nonselective-type
11	34	72.3	493	2	S63060	hypothetical prote
12	34	72.3	493	2	T08407	cathepsin coat asse
13	33	70.2	145	2	AD5334	antibiotic-induced
14	33	70.2	290	2	A23534	protein kinase byr
15	33	70.2	340	1	OKBYR1	lysine acetyltrans
16	33	70.2	391	2	S39816	hypothetical prote
17	32	68.1	102	2	D71042	hypothetical prote
18	32	68.1	177	2	G75013	hypothetical prote
19	32	68.1	190	2	AF1793	glutamine amidotra
20	32	68.1	272	2	AF3515	extragenic suppress
21	32	68.1	346	2	H70722	probable transfera
22	32	68.1	384	1	SYECSM	methionine adenosy
23	32	68.1	384	2	AB0877	S-adenosylmethioni
24	32	68.1	384	2	B91106	methionine adenosy
25	32	68.1	384	2	E85951	methionine adenosy
26	32	68.1	409	2	AC0114	hypothetical prote
27	32	68.1	421	2	C87319	hypothetical prote
28	32	68.1	436	2	E95534	polynucleotide add
29	32	68.1	454	2	G97701	hypothetical prote
	32	68.1	454	2	S58148	hypothetical prote

30	32	68.1	939	2	C70876	hypothetical prote
31	32	68.1	1047	2	S19508	MSH3 protein - Yea
32	31	66.0	83	2	S28120	gas-vesicle operon
33	31	66.0	135	2	E97460	hypothetical prote
34	31	66.0	148	2	S07196	meiosis and sporul
35	31	66.0	158	2	T29432	hypothetical prote
36	31	66.0	233	2	A13495	probable hydrolase
37	31	66.0	246	2	T05961	1,3-beta-glucanase
38	31	66.0	246	2	T05955	1,3-beta-glucanase
39	31	66.0	299	2	C72242	hypothetical prote
40	31	66.0	340	2	UC7695	G protein-coupled
41	31	66.0	391	2	A11701	hypothetical prote
42	31	66.0	396	2	H83576	methionine adenosy
43	31	66.0	421	2	C87545	hypothetical prote
44	31	66.0	561	2	G95367	probable methyl-ac
45	31	66.0	592	2	T25837	hypothetical prote

#### ALIGNMENTS

RESULT 1  
138491  
nucleophosmin/anaplastic lymphoma kinase mutant fusion protein - human  
C/Species: Homo sapiens (man)  
C/Date: 15-Feb-1996 #sequence\_revision 15-Feb-1996 #text\_change 20-Apr-2000  
C/Accession: 138491  
R/Morris, S.W.; Kirstein, M.N.; Valentine, M.B.; Dittmer, K.G.; Shapito, D.N.; Saltman,  
Science 263, 1281-1284, 1994  
A/Title: Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's  
A/Reference number: A53182; M01D:94167588; PMID:8122112  
A/Accession: 138491  
A/Molecule type: mRNA  
A/Residues: 1-680 <MOR>  
A/Cross-references: EMBL:U04946; NID:G609341; PIDN:AA58698.1; PID:G609342  
C/Comment: This sequence is the chimeric product of a translocation mutation.  
C/Genetics:  
A/Gene: NPM1/ALK  
A/Map position: 5/2p23-2p23  
C/Keywords: fusion protein

Query Match 100.0%; Score 47; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDILHV 10  
Db 280 SLAMDILHV 289

RESULT 2  
T30200  
protein-tyrosine kinase (EC 2.7.1.112) alk - mouse  
N/Alternate names: anaplastic lymphoma kinase  
C/Species: Mus musculus (house mouse)  
C/Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T30200  
R/Iwanata, T.; Fujimoto, J.; Wen, D.; Cupples, R.; Bucay, N.; Arakawa, T.; Mori, S.; Ra  
Oncogene 14, 439-449, 1997  
A/Title: Molecular characterization of ALK, a receptor tyrosine kinase expressed specif  
A/Reference number: Z20774; M01D:97118863; PMID:9053841  
A/Accession: T30200  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1621 <TMS>  
A/Cross-references: UNIPROT:P97793; EMBL:DB3002; NID:G1864006; PIDN:BA11673.1; PID:G18  
A/Experimental source: brain and testis  
C/Genetics:  
A/Gene: alk  
C/Function:  
A/Description: may play an important role in development of the brain  
C/Keywords: ATP, phosphotransferase; tyrosine-specific protein kinase

Query Match 100.0%; Score 47; DB 2; Length 1621;  
 Best Local Similarity 100.0%; Pred. No. 0.34;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
 DB 1224 SLAMDLHLV 1233

## RESULT 3

S65091

geranyltransferase type II beta chain - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe

C>Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004

C:Accession: S65091; T37741

R:Godfrey, R.; Davey, J.

Yeast 12, 479-483, 1996

A:Title: Sequence of pGDL1, a gene for the beta subunit of the type-II geranyltransferase

A:Reference number: S65091; MUID:96310627; PMID:8740421

A:Accession: S65091

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-311 <GOD>

A:Cross-references: UNIPROT:P46960; EMBL:X82183; NID:G1033065; PIDN:CAA63094.1; PID:G103

R:Reger, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, January 1999

A:Reference number: 221743

A:Accession: T37741

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-311 <RFE>

A:Cross-references: EMBL:AL035248; PIDN:CAA22847.1; GSPDB:GN00066; SPDB:SPAC167.02

A:Experimental source: strain 972h-; cosmid c167

C:Genetics:

A:Gene: SPAC167.02

A:Map position: 1

A:Introns: 18/3; 69/1; 108/1; 180/1; 224/3

C:superfamily: DPE1 protein

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query Match 76.6%; Score 36; DB 2; Length 1039;  
 Best Local Similarity 77.8%; Pred. No. 39;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 9  
 DB 514 SLAMDLHLV 522

## RESULT 5

T41581

hypothetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T41581

R:Murphy, L.; Harris, D.; Wood, V.; Lyne, M.H.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1998

A:Reference number: 222002

A:Accession: T41581

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4717 <MUR>

A:Cross-references: UNIPROT:Q94248; EMBL:AL031546; PIDN:CAA20864.1; GSPDB:GN00068; SPDB

A:Experimental source: strain 972h(-)

C:Genetics:

A:Gene: SPDB:SPCC737.08

A:Map position: 3

C:Superfamily: mda5in (AAA Arpase with von Willebrand factor type A (vwa) domain)

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

RESULT 7  
 A41591  
 endothelin receptor B precursor - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 09-Jul-2004  
 C:Accession: A41591; A61251; A41184; S63513

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

R,Saito, Y.; Mizuno, T.; Itakura, M.; Suzuki, Y.; Ito, T.; Hagiwara, H.; Hirose, S.  
 J. Biol. Chem. 266, 23433-23437, 1991  
 A>Title: Primary structure of bovine endothelin ET-B receptor and identification of sign  
 A/Reference number: A41591; MUID:92078223; PMID:1660473  
 A/Accession: A41591  
 A/Molecule type: mRNA  
 A/Residues: 1-441 <SAK>  
 A/Cross-references: UNIPROT:P28088; GB:D90456; NID:g217553; PIDN:BA14422.1; PID:g217554  
 R,Hagiwara, H.; Kozuka, M.; Sakaguchi, H.; Eguchi, S.; Ito, T.; Hirose, S.  
 J. Cardiovasc. Pharmacol. 17(Suppl. 7), S117-S118, 1991  
 A>Title: Separation and purification of 34- and 52-kDa species of bovine lung endothelin  
 A/Reference number: A61251; MUID:92219666; PMID:1125302  
 A/Accession: A61251  
 A/Molecule type: protein  
 A/Residues: 304-315/424-432 <HAG>  
 R,Kozuka, M.; Ito, T.; Hirose, S.; Lodhi, K.M.; Hagiwara, H.  
 J. Biol. Chem. 266, 16892-16896, 1991  
 A>Title: Purification and characterization of bovine lung endothelin receptor.  
 A/Reference number: A41184; MUID:91358493; PMID:1653249  
 A/Accession: A41184  
 A/Molecule type: protein  
 A/Residues: 144-127; S, 263-269/304-316/418-421/424-432 <KOZ>  
 R,Hick, S.; Heidemann, T.; Soskic, V.; Mueller-Esterl, W.; Godovac-Zimmermann, J.  
 Eur. J. Biochem. 234, 251-257, 1995  
 A>Title: Isolation of the endothelin B receptor from bovine lung. Structure, signal sequ  
 A/Reference number: S63513; MUID:96096746; PMID:8529649  
 A/Accession: S63513  
 A/Molecule type: protein  
 A/Residues: 1-'X',3-'X',9-10/25-35/36-42/73-78/79-87/284-290/291-296,'X',298 <HIC>  
 C/Superfamily: endothelin receptor B  
 C/Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein  
 F/1-26/Domain: signal sequence #status experimental <SIG>  
 F/27-441/Product: endothelin receptor B #status experimental <MAT>  
 F/101-126/Domain: transmembrane #status predicted <TM1>  
 F/117-162/Domain: transmembrane #status predicted <TM2>  
 F/175-196/Domain: transmembrane #status predicted <TM3>  
 F/218-442/Domain: transmembrane #status predicted <TM4>  
 F/211-295/Domain: transmembrane #status predicted <TM5>  
 F/324-449/Domain: transmembrane #status predicted <TM6>  
 F/362-389/Domain: transmembrane #status predicted <TM7>  
 F/60/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 74.5%; Score 35; DB 2; Length 441;  
 Best Local Similarity 70.0%; Pred. No. 25;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Cy 1 SLAMDLHLV 10  
 Db 141 SLALGDLHI 150

RESULT 8  
 S13425  
 endothelin receptor ETB - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 02-Dec-1993 #sequence revision 10-Nov-1995 #text\_change 09-Jul-2004  
 C/Accession: S13425; A48197  
 R,Sakurai, T.; Yanagisawa, M.; Takawa, Y.; Miyazaki, H.; Kimura, S.; Goto, K.; Masaki, T.  
 Nature 348, 732-735, 1990  
 A>Title: Cloning of a cDNA encoding a non-isopeptide-selective subtype of the endothelin  
 A/Reference number: S13425; MUID:91080924; PMID:2175397  
 A/Accession: S13425  
 A/Molecule type: mRNA  
 A/Residues: 1-441 <SAK>  
 A/Cross-references: UNIPROT:P21451; GB:X57764; NID:G56122; PIDN:CAA40916.1; PID:G56123  
 R,Hori, S.; Komatsu, Y.; Shigemoto, R.; Mizuno, N.; Nakanishi, S.  
 Endocrinology 130, 1885-1895, 1992  
 A>Title: Distinct tissue distribution and cellular localization of two messenger ribonuc  
 A/Reference number: A49197; MUID:92191882; PMID:1112429  
 A/Accession: A49197  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: nucleic acid  
 A/Residues: 1-65; SSAP, 69-441 <HOR>

A/Experimental source: lung  
 A/Note: sequence extracted from NCBI backbone (NCBI:89468)  
 A/Superfamily: endothelin receptor B  
 C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 74.5%; Score 35; DB 2; Length 441;  
 Best Local Similarity 70.0%; Pred. No. 25;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Cy 1 SLAMDLHLV 10  
 Db 141 SLALGDLHI 150

RESULT 9  
 JQ1042  
 endothelin receptor B precursor - human  
 C/Species: Homo sapiens (man)  
 C/Date: 10-Sep-1993 #sequence revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: A46609; JQ1042; JN0072; JQ1082; B45481  
 R,Iwata, H.; Nakao, K.; Takaya, K.; Hosoda, K.; Ogawa, Y.; Nakanishi, S.; Imura, H.  
 J. Biol. Chem. 268, 3463-3470, 1993  
 A>Title: The human endothelin-B receptor gene. Structural organization and chromosomal  
 A/Reference number: A46609; MUID:91155196; PMID:8429023  
 A/Accession: A46609  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-442 <ARA>  
 A/Cross-references: UNIPROT:P24530; GB:D13168; NID:G285924; PIDN:BA02445.1; PID:G285924  
 R,Ogawa, Y.; Nakao, K.; Araki, H.; Nakagawa, O.; Hosoda, K.; Suga, S.; Nakanishi, S.; Imu  
 Biochem. Biophys. Res. Commun. 178, 248-255, 1991  
 A>Title: Molecular cloning of a non-isopeptide-selective human endothelin receptor.  
 A/Reference number: JQ1042; MUID:91298956; PMID:1648908  
 A/Accession: JQ1042  
 A/Molecule type: mRNA  
 A/Residues: 1-442 <OGA>  
 A/Cross-references: GB:D90402; NID:G219651; PIDN:BA14398.1; PID:G219652  
 A/Experimental source: placenta  
 R,Nakamura, M.; Takayanagi, R.; Sakai, Y.; Sakamoto, S.; Hagiwara, H.; Mizuno, T.; Saito  
 Biochem. Biophys. Res. Commun. 177, 34-39, 1991  
 A>Title: Cloning and sequence analysis of a cDNA encoding human non-selective type of en  
 A/Reference number: JN0072; MUID:91254298; PMID:1710450  
 A/Accession: JN0072  
 A/Molecule type: mRNA  
 A/Residues: 1-442 <NAK>  
 A/Cross-references: GB:M74921; NID:G182275; PIDN:AAA58465.1; PID:G182276  
 A/Experimental source: liver  
 R,Sakamoto, A.; Yanagisawa, M.; Sakurai, T.; Takawa, Y.; Yanagisawa, H.; Masaki, T.  
 Biochem. Biophys. Res. Commun. 178, 656-663, 1991  
 A>Title: Cloning and functional expression of human cDNA for the ETB endothelin receptor  
 A/Reference number: JQ1082; MUID:91315496; PMID:1713452  
 A/Accession: JQ1082  
 A/Molecule type: mRNA  
 A/Residues: 1-9,'P',11-442 <SAK>  
 A/Cross-references: GB:S44866; NID:G233233; PIDN:AA19411.1; PID:G233234  
 A/Experimental source: jejunum  
 R,Elsenhoutbagy, N.A.; Korman, D.R.; Wu, H.L.; Sylvester, D.R.; Lee, J.A.; Nuthalaganti, B  
 J. Biol. Chem. 268, 3873-3879, 1993  
 A>Title: Molecular characterization and regulation of the human endothelin receptors.  
 A/Reference number: A45481; MUID:93179382; PMID:8440682  
 A/Accession: A45481  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-442 <ELS>  
 A/Cross-references: GB:L06623; NID:G181958; PIDN:AAA52342.1; PID:G181959  
 A/Experimental source: lung  
 A/Note: sequence extracted from NCBI backbone (NCBI:125754, NCBI:P.125755)  
 C/Comment: This endothelin receptor is non-isopeptide selective.  
 C/Genetics:  
 A/Gene: GDB:EDNRB; ETB  
 A/Cross-references: GDB:129075; OMIM:131244  
 A/Map position: 13q22-13q22  
 C/Superfamily: endothelin receptor B

C/Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein  
 F1-26/Domain: signal sequence #status predicted <SIG>  
 F127-442/Product: endothelin receptor B #status predicted <MAT>  
 F127-442/Product: endothelin receptor B #status predicted <TM1>  
 F139-137/Domain: transmembrane #status predicted <TM2>  
 F139-137/Domain: transmembrane #status predicted <TM2>  
 F176-197/Domain: transmembrane #status predicted <TM3>  
 F219-224/Domain: transmembrane #status predicted <TM4>  
 F219-224/Domain: transmembrane #status predicted <TM5>  
 F325-350/Domain: transmembrane #status predicted <TM6>  
 F363-389/Domain: transmembrane #status predicted <TM7>  
 F159/Binding site: carbohydrate (asn) (covalent) #status predicted  
 F305,413,419,435,436,440/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 74.5% Score 35; DB 1; Length 442;  
 Best Local Similarity 70.0%; Pred. No. 25;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIHV 10  
 |||:||||:  
 Db 142 SLALGDLH 151

## RESULT 10

157950 nonselective-type endothelin receptor - rat  
 C/Species: Rattus sp. (rat)  
 C/Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999

C/Accession: 157950  
 R.Cheng, H.F.; Su Y.M.; Yeh, J.R.; Chang, K.J.  
 Mol. Pharmacol. 44, 533-538, 1993

A/Title: Alternative transcript of the nonselective-type endothelin receptor from rat brain  
 A/Reference number: 157950; MUID:93382424; PMID:8371713

A/Accession: 157950

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-442 <RES>

A/Cross-references: GB:56535; NID:G410692; PIDN:AB28172.1; PID:G410693

C/Superfamily: endothelin receptor B

Query Match 74.5% Score 35; DB 2; Length 442;  
 Best Local Similarity 70.0%; Pred. No. 25;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIHV 10  
 |||:||||:  
 Db 142 SLALGDLH 151

## RESULT 11

S63060 hypothetical protein YNL119w - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein N1913

C/Species: Saccharomyces cerevisiae

C/Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 09-Jul-2004

C/Accession: S63060; S67339

R.De Antoni, A.; D'Angelo, M.; Dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.

submitted to the Protein Sequence Database, April 1996

A/Reference number: S63047

A/Accession: S63060

A/Molecule type: DNA

A/Residues: 1-493 <DEA>

A/Cross-references: UNIPROT:P53923; EMBL:Z71396; NID:G1302046; PID:E239801; PID:G1302048

A/Experimental source: strain S288C

Rid'Antoni, A.; d'Angelo, M.; dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.

submitted to the EMBL Data Library, February 1996

A/Description: The DNA sequence of cosmid 14-13b from chromosome XIV of Saccharomyces cerevisiae

A/Reference number: S67327

A/Accession: S67339

A/Molecule type: DNA

A/Residues: 1-493 <DAN>

A/Cross-references: EMBL:Z69382; NID:G1183941; PID:E221827; PID:G1183954

C/Genetics: SGD:S0005063

A/Map position: 14L  
 A/Note: YNL119w

Query Match 72.3% Score 34; DB 2; Length 493;  
 Best Local Similarity 66.7%; Pred. No. 45;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIHV 9  
 |||:||||:  
 Db 76 SLVMDLHV 84

## RESULT 12

T08407

clathrin coat assembly protein homolog F18B3.140 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C/Accession: T08407

R.Quetier, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salanoubat

submitted to the Protein Sequence Database, May 1999

A/Reference number: T16409

A/Accession: T08407

A/Molecule type: DNA

A/Residues: 1-145 <QUB>

A/Cross-references: UNIPROT:Q9SVL3; EMBL:AL049862; GSPDB:GN00061; ATSP:F18B3.140

A/Experimental source: cultivar Columbia; BAC clone F18B3

C/Genetics:

A/Map position: 3

A/Introns: 23/3; 38/1; 58/3; 91/1; 114/3

C/Superfamily: mouse clathrin-associated protein 19

QY 2 LAMDLIHV 10  
 |||:||||:  
 Db 83 LAMDLIHV 91

## RESULT 13

AD3534 antibiotic-induced protein drp3 (imported) - Brucella melitensis (strain 16M)

C/Species: Brucella melitensis

C/Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004

C/Accession: AD3534

R.DelVecchio, V.G.; Kapral, V.; Redker, R.J.; Patra, G.; Muir, C.; Los, T.; Ivanova,

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A/Reference number: AD3252; PMID:11756688

A/Accession: AD3534

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-230 <KUR>

A/Cross-references: UNIPROT:Q8YDH8; GB:AE008918; PIDN:AAL53439.1; PID:G17984337; GSPDB:G

A/Experimental source: strain 16M

C/Genetics:

A/Map position: 11

Query Match 70.2% Score 33; DB 2; Length 290;  
 Best Local Similarity 66.7%; Pred. No. 42;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIHV 9  
 |||:||||:  
 Db 26 SLAVDLHV 34

## RESULT 14

OKBYR1



protein kinase byr1 (EC 2.7.1.-) - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004

C:Accession: S00473; T38054

R:Nadin-Davies, S.A.; Nadin, A.

EMBO J. 7, 985-993, 1988

A:Title: A gene which encodes a predicted protein kinase can restore some functions of

A:Reference number: S00473; PMID:8826442; PMID:3042386

A:Accession: S00473

A:Molecule type: DNA

A:Residues: 1-340 <NAD>

A:Cross-references: UNIPROT:P10506; EMBL:X07445; NID:G4918; PIDN:CAA30326.1; PID:G4919

R:Lyse, G.; Churcher, C.M.; Barrall, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, February 1995

A:Reference number: 221765

A:Accession: T38054

A:Status: Preliminary; translated from GH/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-340 <LYB>

A:Cross-references: EMBL:Z69239; PIDN:CAA93222.1; GSPDB:GN00066; SPDB:SPAC1D4.13

A:Experimental source: Strain 972-; cosmid CID4

C:Comment: Inactivation of this gene does not interfere with mitotic growth but prevents

C:Genetics:

A:Gene: byr1

A:Map position: 1

C:Superfamily: Kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threonine

F:64.320/Domain: protein kinase homology <KIN>

F:72-80/Region: protein kinase ATP-binding motif

F:93/Active site: Lys #status predicted

Query Match 70.2%; Score 33; DB 1; Length 340;

Best Local Similarity 66.7%; Pred. No. 49;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIH 9

Db 267 SIGILDLIH 275

#### RESULT 15

S39816

lysine acetyltransferase - Yeast (Yarrowia lipolytica)

C:Species: Yarrowia lipolytica; Candida lipolytica

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C:Accession: S39816; S18927

R:Beckerich, J.M.; Lambert, M.; Gallardin, C.

Curr. Genet. 25, 24-29, 1994

A:Title: LYC1 is the structural gene for lysine N-6-acetyl transferase in yeast.

A:Reference number: S39816; MUID:94363731; PMID:8082161

A:Accession: S39816

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-391 <BEC>

A:Cross-references: UNIPROT:P41929; EMBL:X63548; NID:G5246; PIDN:CAA45112.1; PID:G5247

C:Genetics:

A:Gene: lyc1

C:Superfamily: Yarrowia lipolytica lysine acetyltransferase

Query Match 70.2%; Score 33; DB 2; Length 391;

Best Local Similarity 66.7%; Pred. No. 57;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIH 9

Db 166 SLRLMDLIH 174

Search completed: November 30, 2004, 07:41:22  
Job time : 21 secs

**This Page Blank (uspto)**

Tue Nov 30 08:49:20 2004

us-10-008-377a-1.open.rup

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12; Search time 76.4103 Seconds  
(without alignments)  
75.301 Million cell updates/sec

Title: US-10-008-377a-1  
Perfect score: 47  
Sequence: 1 SLAMDLMHV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: 1: uniprot\_sprot;\*  
2: uniprot\_trembl;\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	351	2 Q8CJH1	Q8CJH1 ratius nov
2	47	100.0	803	2 Q8TDJ5	Q8TDJ5 homo sapien
3	47	100.0	1620	1 ALK_HUMAN	Q9UM73 homo sapien
4	47	100.0	1621	1 ALK_MOUSE	P97793 mus musculu
5	37	78.7	311	1 PTBL_SCHPO	P46960 schizosach
6	37	78.7	792	2 Q8ECI3	Q8ECI3 shewanelia
7	37	78.7	1161	2 Q8UFV0	Q8UFV0 brachydanto
8	36	76.6	376	2 Q9XHX1	Q9XHX1 cryza sativ
9	36	76.6	461	2 Q7SBK0	Q7SBK0 asblya goss
10	36	76.6	461	2 AAS51497	AAS51497 asblya go
11	36	76.6	763	2 Q7S4B6	Q7S4B6 neuropsora
12	36	76.6	872	2 Q8RJM3	Q8RJM3 xanthomonas
13	36	76.6	989	2 Q23723	Q23723 arabidopsis
14	36	76.6	1039	2 Q23461	Q23461 arabidopsis
15	36	76.6	1065	2 Q7Q8Y1	Q7Q8Y1 glardia lam
16	36	76.6	4717	2 Q94248	Q94248 schizosach
17	35	74.5	152	2 Q9F1S6	Q9F1S6 streptococc
18	35	74.5	208	2 Q9F1S3	Q9F1S3 streptococc
19	35	74.5	232	2 Q29887	Q29887 archaeglob
20	35	74.5	276	2 Q9F1T3	Q9F1T3 streptococc
21	35	74.5	276	2 Q93R75	Q93R75 streptococc
22	35	74.5	312	2 Q8AJE1	Q8AJE1 mus musculu
23	35	74.5	347	1 ETRR_COTJA	Q90328 coturnix co
24	35	74.5	356	2 Q8UHV4	Q8UHV4 gallus gall
25	35	74.5	382	2 Q7NRP3	Q7NRP3 chromobacte
26	35	74.5	426	1 ETRR_CANFA	P56497 canis famli
27	35	74.5	426	1 Q6MX81	Q6MX81 azoarcus sp
28	35	74.5	426	2 Q9DGM2	Q9DGM2 brachydanto
29	35	74.5	426	2 CAF21977	CAF21977 azoarcus
30	35	74.5	434	2 Q9MYZ5	Q9MYZ5 canis famli
31	35	74.5	441	1 ETRR_BOVIN	P28088 bos taurus

32	35	74.5	441	2 Q9N0W7	Q9N0W7 oryctolagus
33	35	74.5	442	1 ETRR_HUMAN	P24530 homo sapien
34	35	74.5	442	1 ETRR_MOUSE	P40302 mus musculu
35	35	74.5	442	1 ETRR_RAT	P21451 rattus norv
36	35	74.5	442	2 Q9UD23	Q9UD23 homo sapien
37	35	74.5	442	2 Q8KIV6	Q8KIV6 mus musculu
38	35	74.5	442	2 AAS39516	AAS39516 homo sapi
39	35	74.5	442	2 BAC36337	BAC36337 mus muscu
40	35	74.5	442	2 BAC38409	BAC38409 mus muscu
41	35	74.5	442	2 BAC38908	BAC38908 mus muscu
42	35	74.5	443	2 BAC39465	BAC39465 mus muscu
43	35	74.5	443	1 ETRR_HORSE	O62709 equus cabal
44	35	74.5	443	1 ETRR_PIG	P35463 sus scrofa
45	35	74.5	612	2 Q8VXF6	Q8VXF6 lupinus lut

#### ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT;	351 AA.
ID Q8CJH1			
AC Q8CJH1			
DT 01-MAR-2003 (TEMBLrel. 23, Created)			
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)			
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)			
DE Anaplastic lymphoma kinase (Fragment).			
GN Name=alk;			
OS Rattus norvegicus (Rat).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX NCBI_TaxID=10116;			
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
RL EMBL: AB073169; BAC21663.1; -			
DR GO: GO:0005224; F1ATP binding; IEA.			
DR GO: GO:0004713; F1protein-tyrosine kinase activity; IEA.			
DR GO: GO:0016740; F1transferase activity; IEA.			
DR GO: GO:0006468; F1protein amino acid phosphorylation; IEA.			
DR InterPro: IPR011009; Kinase, like.			
DR InterPro: IPR000719; Prot kinase.			
DR InterPro: IPR002011; RecepttyrkinaII.			
DR InterPro: IPR001245; Tyr kinase.			
DR InterPro: IPR008266; Tyr_kinase_AS.			
DR Pfam: PF00069; PKinase; 1.			
DR PRINTS: PR0109; TYRKINASE.			
DR PRODOM: PD000001; Prot kinase; 1.			
DR SMART: SM00219; TyKc; 1.			
DR PROSITE: PS00107; PROTEIN KINASE ATP; UNKNOWN_1.			
DR PROSITE: PS00111; PROTEIN KINASE DOM; 1.			
DR PROSITE: PS00109; PROTEIN KINASE TYR; 1.			
DR PROSITE: PS00239; RECEPTOR TYR_KIN II; UNKNOWN_1.			
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.			
FT NON_TER			
FT TER			
FT NON_TER			
FT TER			
SEQUENCE	351 AA; 38984 MW; 6ADCEG220C76B125 CRC64;		
Query Match	100.0%; Score 47; DB 2; Length 351;		
Best Local Similarity	100.0%; Pred. No. 0.41;		
Matches	10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1 SLAMDLMHV 10		
DB	223 SLAMDLMHV 232		
RESULT 2			
Q8TDJ5	PRELIMINARY;	PRT;	803 AA.
ID Q8TDJ5			

```

AC Q8TJ5;
DT 01-JUN-2002 (TRENBLREL. 21, Created)
DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE TRK-fused gene/anaplastic large cell lymphoma kinase extra long
DE form.
GN Name=TRK/ALK fusion;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21940189; PubMed=1194373;
RA Hernandez L., Bea S., Bellosillo B., Pinyol M., Falini B., Ott G.,
RA Pulford K., Rosenwald A., Morris S.W., Fernandez A., Santos E.,
RA Campo E.;
RT "Diversity of genomic breakpoints in TRK-ALK translocations in
RT anaplastic large cell lymphomas: identification of a new TRK-ALK(XL)
RT chimeric gene with transforming activity.";
RL Am. J. Pathol. 160:1487-1494(2002).
DR EMBL: AF390893; AAM17922.1; -
DR HSSP: O62838; 1LUF.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO: GO:0018740; F:transferase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000270; CDR P1.
DR InterPro: IPR000719; Prot. kinase.
DR InterPro: IPR002011; Recepttyrkinasit.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00564; P1; 1.
DR Pfam: PF00069; Pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Prot. kinase; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKOWN_1.
DR PROSITE: PS50019; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00239; RECEPTOR_TYR_KIN_11; UNKOWN_1.
DR ATP-binding; Kinase; transferase; Tyrosine-protein kinase.
KW SEQUENCE 803 AA; 88670 MW; 890826D49EBE9EF CRC64;
SQ
Query Match 100.0%; Score 47; DB 2; Length 803;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLAMF1LHV 10
DB 403 SLAMF1LHV 412

```

```

RA Morris S.W., Naeve C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,
RA Witte D.P.;
RT "ALK, the chromosome 2 gene locus altered by the t(2;5) in non-
RT Hodgkin's lymphoma, encodes a novel neutral receptor tyrosine kinase
RT that is highly related to leukocyte tyrosine kinase (LTK).";
RL Oncogene 14:2175-2186(1997).
RN [2]
RP ERRATUM.
RA Morris S.W., Naeve C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,
RA Witte D.P.;
RL Oncogene 15:2883-2883(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97178863; PubMed=9053841;
RA Iwahara T., Fujimoto J., Wen D., Cupples R., Bucay N., Arakawa T.,
RA Mori S., Ratzkin B., Yamamoto T.;
RT "Molecular characterization of ALK, a receptor tyrosine kinase
RT expressed specifically in the nervous system.";
RL Oncogene 14:439-449(1997).
RN [4]
RP PARTIAL SEQUENCE FROM N.A., AND CHROMOSOMAL TRANSLOCATION.
RX MEDLINE=94167588; PubMed=812212;
RA Morris S.W., Kirstein M.N., Valentine M.B., Dittmer K.G.,
RA Shapiro D.N., Saltman D.L., Look A.T.;
RT "Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in
RT non-Hodgkin's lymphoma.";
RL Science 263:1281-1284(1994).
CC -1- FUNCTION: Orphan receptor with a tyrosine-protein kinase activity.
CC Appears to play an important role in the normal development and
CC function of the nervous system.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in brain and CNS. Also expressed in
CC the small intestine and testis, but not in normal lymphoid cells.
CC -1- PTM: N-glycosylated.
CC -1- DISEASE: A form of non-Hodgkin's lymphoma is characterized by a
CC chromosomal translocation t(2;5) (p23;q35) that involves NPM1 and
CC ALK.
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
CC receptor subfamily.
CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -1- SIMILARITY: Contains 2 WAM domains.
CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW=http://www.infobio.genet.fr/services/chromocancer/genes/ALK.htm1".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
DR EMBL: U62540; AAB71619.1; -
DR EMBL: U66559; AAC51104.1; -
DR HSSP: O62838; 1LUF.
DR Genew: HGNC:427; ALK.
DR MIM: 105590; -
DR GO: GO:0016021; C:integral to membrane; NAS.
DR GO: GO:0007399; P:neurogenesis; NAS.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR002172; LDL_receptor_A.
DR InterPro: IPR000998; WAM.
DR InterPro: IPR000719; Prot. kinase.
DR InterPro: IPR002011; Recepttyrkinasit.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00629; WAM; 1.
DR Pfam: PF00069; Pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Prot. kinase; 1.
DR SMART: SM00192; LDLa; 1.

```

DR SMART; SM00219; TykC; 1.  
 DR PROSITE; PS01209; LDLR\_1; FALSE NEG.  
 DR PROSITE; PS00668; LDLR\_2; FALSE NEG.  
 DR PROSITE; PS00740; MAM\_1; FALSE NEG.  
 DR PROSITE; PS00660; MAM\_2; 2.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 DR ATP-binding; Chromosomal translocation; Phosphorylation;  
 KW ATP-binding; Receptor; Repeat; Signal; Transferrase; Transmembrane;  
 KW Tyrosine-protein kinase.  
 FT SIGNAL 1 18  
 FT CHAIN 19 1620 Potential.  
 FT DOMAIN 19 1038 ALK tyrosine kinase receptor.  
 FT TRANSMEM 1039 1059 Extracellular (Potential).  
 FT DOMAIN 1060 1620 Potential.  
 FT DOMAIN 1060 1620 Cytoplasmic (Potential).  
 FT DOMAIN 264 427 MAM 1.  
 FT DOMAIN 437 473 LDL-receptor class A.  
 FT DOMAIN 478 636 MAM 2.  
 FT DOMAIN 1116 1392 Protein kinase.  
 FT NP\_BIND 816 940 Gly-rich.  
 FT BINDING 1122 1130 ATP (By similarity).  
 FT ACT\_SITE 1150 1150 ATP (By similarity).  
 FT MOD\_RES 1249 1249 Proton acceptor (By similarity).  
 FT MOD\_RES 1282 1282 Phosphotyrosine (by autocatalysis) (By similarity).  
 FT CARBOHYD 169 169 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 244 244 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 285 285 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 324 324 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 411 411 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 424 424 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 445 445 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 563 563 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 571 571 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 627 627 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 627 627 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 709 709 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 808 808 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 863 863 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 864 864 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 886 886 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 986 986 N-linked (GlcNAc...) (Potential).  
 FT CONFLICT 36 36 S -> P (in Ref. 3).  
 FT CONFLICT 1491 1491 K -> R (in Ref. 3).  
 FT CONFLICT 1529 1529 D -> E (in Ref. 3).  
 SQ SEQUENCE 1620 AA; 176417 MW; A62604B242961E1E CRC64;

Query Match 100.0%; Score 47; DB 1; Length 1620;  
 Best Local Similarity 100.0%; Pred. No. 2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLAMF1LHV 10  
 DB 1220 SLAMF1LHV 1229

RESULT 4  
 ID ALK\_MOUSE STANDARD; PRT; 1621 AA.  
 AC P97793;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE ALK tyrosine kinase receptor precursor (EC 2.7.1.112) (Anaplastic  
 DE lymphoma kinase).  
 GN Name=Alk;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Brain, and Testis;  
 RX MEDLINE=9717863; PubMed=9053841;  
 RA Iwihara T., Fujimoto J., Wen D., Cupples R., Bucay N., Arakawa T.,  
 RA Mori S., Ratzkin B., Yamamoto T.;  
 RT "Molecular characterization of ALK, a receptor tyrosine kinase  
 RT expressed specifically in the nervous system";  
 RL Oncogene 14:439-449(1997).  
 CC -1 FUNCTION: Orphan receptor with a tyrosine-protein kinase activity.  
 CC Appears to play an important role in the normal development and  
 CC function of the nervous system.  
 CC -1 CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein (Probable).  
 CC -1 SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin  
 CC receptor subfamily.  
 CC -1 SIMILARITY: Contains 2 MAM domains.  
 CC -1 SIMILARITY: Contains 2 MAM domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC EMBL; D83002; BAA11673.1; -  
 DR PIR; T30200; T30200.  
 DR HSSP; P08069; 1JQH.  
 DR MGD; MGI:103305; Alk.  
 DR GO; GO:0005515; F-protein binding; IPI.  
 DR InterPro; IPR011009; Kinase\_like.  
 DR InterPro; IPR002172; LDL\_receptor\_A.  
 DR InterPro; IPR000998; MAM.  
 DR InterPro; IPR00719; Prot\_kinase.  
 DR InterPro; IPR002011; ReceptTyKinsII.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00629; MAM\_1.  
 DR Pfam; PF00629; PKinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00192; LDLR\_1.  
 DR SMART; SM00219; TykC; 1.  
 DR PROSITE; PS01209; LDLR\_1; FALSE NEG.  
 DR PROSITE; PS00668; LDLR\_2; FALSE NEG.  
 DR PROSITE; PS00740; MAM\_1; FALSE NEG.  
 DR PROSITE; PS00660; MAM\_2; 2.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 KW ATP-binding; Chromosomal translocation; Phosphorylation; Receptor; Repeat; Signal;  
 KW Tyrosine-protein kinase.  
 FT SIGNAL 1 18 Potential.  
 FT CHAIN 19 1621 ALK tyrosine kinase receptor.  
 FT DOMAIN 19 1042 Extracellular (Potential).  
 FT TRANSMEM 1043 1063 Potential.  
 FT DOMAIN 1064 1621 Cytoplasmic (Potential).  
 FT DOMAIN 268 431 MAM 1.  
 FT DOMAIN 441 477 LDL-receptor class A.  
 FT DOMAIN 482 640 MAM 2.  
 FT DOMAIN 820 1396 Protein kinase.  
 FT NP\_BIND 1126 1134 Gly-rich.  
 FT BINDING 1154 1154 ATP (By similarity).  
 FT ACT\_SITE 1253 1253 Proton acceptor (By similarity).  
 FT MOD\_RES 1286 1286 Phosphotyrosine (by autocatalysis) (By similarity).  
 FT CARBOHYD 174 174 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 248 248 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 289 289 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 328 328 N-linked (GlcNAc...) (Potential).

```

FT CAROHD 415 415 N-linked (GlcNAc. . .) (Potential).
FT CAROHD 428 428 N-linked (GlcNAc. . .) (Potential).
FT CAROHD 449 449 N-linked (GlcNAc. . .) (Potential).
FT CAROHD 567 567 N-linked (GlcNAc. . .) (Potential).
FT CAROHD 575 575 N-linked (GlcNAc. . .) (Potential).
FT CAROHD 631 631 N-linked (GlcNAc. . .) (Potential).
FT CAROHD 673 673 N-linked (GlcNAc. . .) (Potential).
FT CAROHD 713 713 N-linked (GlcNAc. . .) (Potential).
FT CAROHD 812 812 N-linked (GlcNAc. . .) (Potential).
FT CAROHD 868 868 N-linked (GlcNAc. . .) (Potential).
FT CAROHD 890 890 N-linked (GlcNAc. . .) (Potential).
FT CAROHD 990 990 N-linked (GlcNAc. . .) (Potential).
SQ SEQUENCE 1621 AA; 174919 MW; 16E52BF21AABD22 CRC64;

Query Match 100.0%; Score 47; DB 1; Length 1621;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLAMDLIHV 10
Db 1224 SLAMDLIHV 1233

RESULT 5
PRT1 SCHPO STANDARD; PRT; 311 AA.
ID PRT1 SCHPO STANDARD; PRT; 311 AA.
AC P46360;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE (Type II protein geranyl-geranyltransferase beta subunit) (EC 2.5.1.60)
DE (Type II protein geranyl-geranyltransferase beta subunit) (CGHase-II-
DE beta).
GN Name=ptb1; ORFNames=SPAC167.02;
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96310627; PubMed=8740421;
RT "Sequence of ptb1, a gene for the beta subunit of the type-II
RT geranylgeranyltransferase from the fission yeast Schizosaccharomycetes
RT pombe."
RL Yeast 12:479-482(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21849401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream A.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagsis K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Nisbetic D., Odeli C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,
RA Welfjens I., Vansireels E., Rieger M., Schefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu B., Dreano S., Gloux S., Lelaire V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

```

```

RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Spakovski G.V., Ussery D., Barrett B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -1- FUNCTION: Catalyzes the transfer of a geranyl-geranyl moiety from
CC geranyl-geranyl pyrophosphate to proteins having the C-terminal
CC -XCC or -XCCX, where both cysteines may become modified.
CC -1- CATALYTIC ACTIVITY: 2 geranyl-geranyl diphosphate + protein-
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit.
CC -1- SIMILARITY: Belongs to the protein prenyltransferase beta subunit
CC family.
CC -1- SIMILARITY: Contains 5 PFTB repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: X92183; CA63094.1; -
DR EMBL: A035248; CA22847.1; -
DR PIR: S65091; S65091.
DR HSBP; Q08603; LITX.
DR GeneDB_Spombe; SPAC167.02; -.
DR InterPro; IPR001330; Prenyltrans.
DR InterPro; IPR008930; Terp_Cyc_Toroid.
DR Pfam; PF00432; Prenyltrans; 5.
DR KW: Prenyltransferase; Repeat; Transferase; Zinc.
FT REPEAT 54 95 PFTB 1.
FT REPEAT 102 143 PFTB 2.
FT REPEAT 150 191 PFTB 3.
FT REPEAT 197 239 PFTB 4.
FT REPEAT 246 288 PFTB 5.
FT METAL 224 224 Zinc (By similarity).
FT METAL 226 226 Zinc (By similarity).
FT METAL 276 276 Zinc (By similarity).
SQ SEQUENCE 311 AA; 35092 MW; 2AB617FC769D3B08 CRC64;

Query Match 78.7%; Score 37; DB 1; Length 311;
Best Local Similarity 88.9%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LAMDLIHV 10
Db 91 LAMDLIHV 99

RESULT 6
Q8BC13 PRELIMINARY; PRT; 792 AA.
ID Q8BC13
AC Q8BC13
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein S03159.
DE OrderedLocustNames=S03159;
GN Bacteroides thetaiotaomicens.
OS Bacteroidetes; Bacteroidia; Gammaproteobacteria; Alteromonadales;
OC Bacteriia; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadales; Shevanelia.
CX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NR-1;
RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.B., Gaidos E.J., Nelson W.C.,
RA Reed T.D., Eisen J.A., Seshadri R., Ward N.J., Methe B.A.,
RA Clayton R.A., Meyer J., Teapin A., Scott J., Beaman M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,

```

RA White O., Wolf A.M., Vamathavan J.J., Weidman J.F., Impraim M.,  
 RA Lee K., Berry K.J., Lee C., Mueller J., Knout H.M., Gill J.,  
 RA Ueberback T.R., McDonald L.A., Feldblum T.V., Smith H.O.,  
 RA Venter J.C., Neilson K.H., Fraser C.M.,  
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium  
 RT Shewanella oneidensis",  
 RL Nat. Biotechnol. 20:1118-1123(2002).  
 DR EMBL: AE015753; AAN56159.1; -.  
 DR TIGR: SC3159; -.  
 DR InterPro: IPR010344; DUF940.  
 DR Pfam: PF06082; DUF940; 1.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 792 AA; 88982 MW; F93CEFA5D24966B CRC64;

Query Match 76.7%; Score 37; DB 2; Length 792;  
 Best Local Similarity 80.0%; Pred. No. 1,1e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SLAMDLIHV 10  
 Db 18 SLAMDLIHV 27

RESULT 7  
 Q8UFV0 PRELIMINARY; PRT; 1161 AA.  
 AC Q8UFV0;  
 DT 01-OCT-2002 (TRENBLREL. 22, Created)  
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)  
 DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)  
 DE SI:42107016.1 (Novel protein similar to vertebrate anaplastic lymphoma  
 kinase (ALK) and leukocyte tyrosine kinase receptor precursor (LTK or  
 TYK) [Fragment].  
 CN Name=SI:42107016.1;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Ctenista; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hammett S.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL596022; C943463.1; -.  
 DR HSP; Q62838; IUDF.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO: GO:0004872; F:receptor activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase like  
 DR InterPro: IPR002172; LDI\_receptor\_A.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002011; ReceptTyKinsit.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase\_AS.  
 DR Pfam: PF00057; Ldi\_recept\_a; 1.  
 DR Pfam: PF00069; Pkinase; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR PRODOM: PD000001; Prot\_kinase; 1.  
 DR SMART; SM00192; LDIa; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS01209; LDIa\_1; UNKNOWN\_1.  
 DR PROSITE; PS00068; LDIa\_2; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00039; RECEPTOR\_TYR\_KIN\_II; UNKNOWN\_1.  
 KW ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.  
 FT NON\_TER  
 SQ SEQUENCE 1161 AA; 124851 MW; 601210E708E9AA41 CRC64;

Query Match 78.7%; Score 37; DB 2; Length 1161;

Best Local Similarity 70.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLAMDLIHV 10  
 Db 670 SLAMDLIHV 679

RESULT 8  
 Q9XHX1 PRELIMINARY; PRT; 376 AA.  
 ID Q9XHX1;  
 AC Q9XHX1;  
 DT 01-NOV-1999 (TRENBLREL. 12, Created)  
 DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)  
 DT 01-OCT-2004 (TRENBLREL. 28, Last annotation update)  
 DE Putative geranylgeranyl pyrophosphate synthase.  
 CN Name=OSJNB0049820.22; Synonyms=PG013C11.25;  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriatridae; Oryzaceae; Oryza.  
 OC NCBI\_TaxID=3947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Buell R., Bentio M.-I., Lin X., Mason T.M., Unayam L., Shea T.P.,  
 RA Fujii C.Y., Shen M., Fraser C.M.,  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
 RA Wu J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
 RA Hosokawa S., Masuoka M., Arikawa K., Chiden Y., Hayashi M.,  
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
 RA Hijishita S., Honda M., Ichikawa Y., Idenuma A., Iijima M., Ikeda M.,  
 RA Ikono M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,  
 RA Karasawa M., Katagiri S., Kikuta A., Kodayashi N., Kono T.,  
 RA Machida K., Maehara T., Mizuno H., Mizudayashi T., Nakai Y.,  
 RA Nagasaki N., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terashima K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshitake R., Yukawa K.,  
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.;  
 RT "The genome sequence and structure of rice chromosome 1.";  
 RL Nature 420:312-316(2002).  
 DR EMBL: AC007789; AAD38295.1; -.  
 DR EMBL: AC002865; BAB18334.1; -.  
 DR Gramene; Q9XHX1; -.  
 DR GO: GO:0008299; P:isoprenoid biosynthesis; IEA.  
 DR InterPro: IPR000092; Polyisoprenyl\_synth.  
 DR InterPro: IPR008949; Terpenoid\_synth.  
 DR Pfam: PF00348; Polyisoprenyl\_synth\_1.  
 DR PROSITE; PS00723; POLYISOPRENYL\_SYNTHET\_1; 1.  
 DR PROSITE; PS00444; POLYISOPRENYL\_SYNTHET\_2; 1.  
 DR PROSITE; PS00444; POLYISOPRENYL\_SYNTHET\_2; 1.  
 SQ SEQUENCE 376 AA; 39575 MW; F737BFB8B8B2B426 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 376;  
 Best Local Similarity 70.0%; Pred. No. 84;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SLAMDLIHV 10  
 Db 250 SLAMDLIHV 259

RESULT 9  
 ID Q7SBKO PRELIMINARY; PRT; 461 AA.  
 AC Q7SBKO;  
 DT 05-JUN-2004 (TRENBLREL. 27, Created)  
 DT 05-JUN-2004 (TRENBLREL. 27, Last sequence update)  
 DT 05-JUN-2004 (TRENBLREL. 27, Last annotation update)  
 DE ACR271CP.

```

GN Name=ACR271C; (Yeast) (Eremothecium gossypii).
OS Ashbya gossypii (Yeast); Ascomycota; Saccharomycotina; Saccharomycetes;
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RX PubMed=15001715;
RA Dietrich F.S., Voegelé S., Brachat S., Lerch A., Gates K., Steiner S.,
  Mohr C., Pohlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
  Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
  Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
DR EMBL; AE016888; AAS51497.1; -.
DR AGD; ACR271C; -.
SQ SEQUENCE 461 AA; 52404 MW; 5D6FDBF10928C846 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 461;
Best Local Similarity 88.9%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 80 SLAMDILN 88

RESULT 10
AAS51497 PRELIMINARY; PRT; 461 AA.
ID AAS51497;
AC AAS51497;
DT 23-APR-2004 (TREMBLrel. 27, Created)
DT 23-APR-2004 (TREMBLrel. 27, Last sequence update)
DE ACR271Cp.
GN ACR271C.
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RX PubMed=15001715;
RA Dietrich F.S., Voegelé S., Brachat S., Lerch A., Gates K., Steiner S.,
  Mohr C., Pohlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
  Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
  Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
DR EMBL; AE016888; AAS51497.1; -.
DR EMBL; AE016888; AAS51497.1; -.
SQ SEQUENCE 461 AA; 52404 MW; 5D6FDBF10928C846 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 461;
Best Local Similarity 88.9%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 80 SLAMDILN 88

RESULT 11
07S4B6 PRELIMINARY; PRT; 763 AA.
ID 07S4B6;
AC 07S4B6;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU08177.1;
OS Neurospora crassa.

```

```

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
  Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Reiman B.,
  Eilkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
  Qui D., Iankilev P., Pedersen D., Nelson M., Mashburne M.,
  Saittenuikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
  Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
  Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
  Kamil W., Kamyshelev M., Maccelli E., Blake C., Rudd S., Frisman D.,
  Kyrstova S., Rasmussen C., Metzberg R.L., Perkins D.V., Kroken S.,
  Cogoni C., Macho G., Catchside D., Li W., Platt R.V., Osmani S.A.,
  Desouza C.C., Glaes L., Orbach M.J., Berglund J., Voelker R.,
  Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
  Natvig D.O., Alex L.A., Manhaupt G., Ebbole D.U., Freitag M.,
  Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 010-012003).
CC -!- CAUTION: The sequence shown here is derived from an
  EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
  preliminary data.
CC EMBL; AABX0100370; EAA30334.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR007219; F:Protein_kinase.
DR InterPro; IPR008271; Ser_Chr_kin_AS.
DR Pfam; PF00659; Pkinase; 1.
DR PROSITE; PS00001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 763 AA; 84728 MW; 414B0DE9059D24D5 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 763;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 642 ALAIDLIN 650

RESULT 12
08PUW3 PRELIMINARY; PRT; 872 AA.
ID 08PUW3;
AC 08PUW3;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE ATP-dependent DNA ligase.
GN Name=lig3; Ordered locus names=XAC2414;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferris U.A., Reinach F.C., Farah C.S., Furlan L.R.,
  Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
  Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
  Amarote G., Camaral A.M., Cardoso J., Chamberg F., Cipolina L.P.,
  Ciccarelli R.M.B., Coutinho L.V., Curisio-Santos J.R., El-Dorri H.,
  Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Guber A.,
  Formighieri E.F., Franco M.C., Greggio C.C., Gunter A.,
  Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,

```



RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.N., White F.F.,  
 RA Seubal J.C., Kitajima J.P.,  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities".  
 RL Nature 417:459-463(2002).  
 CC -1- CATALYTIC ACTIVITY: ATP + (deoxyribonucleotide) (N) +  
 CC {deoxyribonucleotide} (M) = AMP + diphosphate +  
 CC {deoxyribonucleotide} (N+M).  
 CC -1- SIMILARITY: Belongs to the ATP-dependent DNA ligase family.  
 DR EMBL: AEO1878; AAM37266.1; -  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003910; F:DNA ligase (ATP) activity; IEA.  
 DR GO: GO:0016874; F:ligase activity; IEA.  
 DR GO: GO:0003676; F:nucleic acid binding; IEA.  
 DR GO: GO:0006310; P:DNA recombination; IEA.  
 DR GO: GO:0006281; P:DNA repair; IEA.  
 DR GO: GO:0006260; P:DNA replication; IEA.  
 DR InterPro: IPR000977; Nucleic acid.  
 DR InterPro: IPR008994; Nucleic acid.  
 DR Pfam: PF04679; DNA\_ligase\_A\_C; 1.  
 DR Pfam: PF01068; DNA\_ligase\_A\_M; 1.  
 DR PROSITE: PS00697; DNA\_LIGASE\_A1; UNKNOWN\_1.  
 DR PROSITE: PS01650; DNA\_LIGASE\_A3; 1.  
 DR ATP-binding; Complete proteome; DNA recombination; DNA repair;  
 KM DNA replication; ligase.  
 SQ SEQUENCE 872 AA; 96305 MW; 3C5AC9FBA2F08921 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 872;  
 Best Local Similarity 70.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIH 10  
 Db 325 ALAVFDLHV 334

RESULT 13

ID 023723 PRELIMINARY; PRT; 989 AA.  
 AC 023723;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative arabinose kinase.  
 GN Name=ISA1;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC euroids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98192529; PubMed=9524266;  
 RA Cy 1., Aubourg S., Sheron S., Cobbett C.S., Cheron A., Kreis M.,  
 RA Lecharny A.,  
 RT "Analysis of a 14-kb fragment containing a putative cell wall gene and  
 RT a candidate for the ARA1, arabinose kinase, gene from chromosome IV of  
 RT Arabidopsis thaliana".  
 RL Gene 209:201-210(1998).  
 DR EMBL: Y14404; CAA74753.1; -  
 DR GO: GO:0005737; C:cytoplasm; IEA.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0016301; F:kinase activity; IEA.  
 DR GO: GO:0016773; F:phosphotransferase activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0008152; P:metabolism; IEA.  
 DR GO: GO:0016310; P:phosphorylation; IEA.  
 DR InterPro: IPR006204; GHMP\_kinase.

DR InterPro: IPR006206; Mey\_galkinase.  
 DR Pfam: PF00288; GHMP\_kinases; 1.  
 DR PRINTS: PRO0959; MEVGALKINASE.  
 KM ATP-binding; Kinase; Transferase.  
 SQ SEQUENCE 989 AA; 10846 MW; CC954177C0888438B CRC64;

Query Match 76.6%; Score 36; DB 2; Length 989;  
 Best Local Similarity 77.8%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIH 9  
 Db 514 SLAMDLIH 522

RESULT 14

ID 023461 PRELIMINARY; PRT; 1039 AA.  
 AC 023461;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Galactokinase like protein.  
 GN Name=d14105w; Synonyms=ATG16130;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC euroids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Stiekema W., Murphy G., Mambutt R., Pohl T., Terry N.,  
 RA Kreis M., Kavanagh T., Entian K.D., Rieger W., James R.,  
 RA Pulgadennech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,  
 RA Jones U., Palme K., Ansgore W., Delsen M., Bancroft I., Mewes H.W.,  
 RA Schueller C., Chalwatzis N.,  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z97340; CAB10392.1; -  
 DR EMBL: AL161543; CAB8655.1; -  
 DR PIR: F71427; F71427.  
 DR GO: GO:0005737; C:cytoplasm; IEA.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0016301; F:kinase activity; IEA.  
 DR GO: GO:0016773; F:phosphotransferase activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0008152; P:metabolism; IEA.  
 DR GO: GO:0016310; P:phosphorylation; IEA.  
 DR InterPro: IPR006204; GHMP\_kinase.  
 DR InterPro: IPR006206; Mey\_galkinase.  
 DR Pfam: PF00288; GHMP\_kinases; 1.  
 DR PRINTS: PRO0959; MEVGALKINASE.  
 KM ATP-binding; Kinase; Transferase.  
 SQ SEQUENCE 1039 AA; 114260 MW; 82843127FCC8665B CRC64;

Query Match 76.6%; Score 36; DB 2; Length 1039;  
 Best Local Similarity 77.8%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIH 9  
 Db 514 SLAMDLIH 522

RESULT 15

Q7GRY1

```

ID Q7ORY1 PRELIMINARY; PRT; 1066 AA.
AC Q7ORY1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GIP_549_14267.17467
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.O., Sogin M.L.; Giardia lamblia genome.
RT "Draft sequence of the Giardia lamblia genome."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL, AACB0100127; EAA37782.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:Protein kinase activity; IEA.
DR GO; GO:0016740; F:Transferase activity; IEA.
DR GO; GO:0006468; P:Protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_Kinase.
DR Pfam; PF00069; Kinase; 1.
DR ProDom; PD000001; Prot_Kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
SQ SEQUENCE 1066 AA; 11953 MW; E214831E6780AE27 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 1066;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

CY 2 LAMDDLHV 10
   :|||||
Db 236 LGLDDLHV 244

```

Search completed: November 30, 2004, 07:17:26  
 Job time : 78.4103 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 77.1795 Seconds

(without alignments)  
46,480 Million cell updates/sec

Title: US-10-008-377A-1

Perfect score: 47

Sequence: 1 SLAMDLLHV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 35672929 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_23sep04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	10	AA022983	AA022983 Human p28
2	47	100.0	45	AB073862	AB073862 Human ALK
3	47	100.0	45	AB074476	AB074476 Tyrosine
4	47	100.0	525	AA078521	AA078521 NPM/ALK f
5	47	100.0	525	AB004368	AB004368 Human exp
6	47	100.0	572	AA078520	AA078520 Human exp
7	47	100.0	675	AB020950	AB020950 Partial A
8	47	100.0	680	AA073858	AA073858 Novel hum
9	47	100.0	680	AB073858	AB073858 Human NPM
10	47	100.0	680	AB095067	AB095067 Human tpa
11	47	100.0	680	AB004347	AB004347 Human exp
12	47	100.0	680	AB004362	AB004362 Human exp
13	47	100.0	680	AB004358	AB004358 Human exp
14	47	100.0	680	AB074472	AB074472 NPM/ALK f
15	47	100.0	1620	AA018498	AA018498 Human ins
16	47	100.0	1620	AB074471	AB074471 Human ins
17	43	91.5	9	AA022985	AA022985 Human p28
18	39	83.0	9	AA022986	AA022986 Human p28
19	36	76.6	376	ADM98606	ADM98606 Human p28
20	36	76.6	376	ADM98621	ADM98621 Geranylge
21	36	76.6	376	ADM98621	ADM98621 Geranylge
22	36	76.6	1039	AB092855	AB092855 Geranylge
23	35	74.5	187	AB070368	AB070368 Novel hum
24	35	74.5	302	AB025901	AB025901 Aspergill
25	35	74.5	334	AB025901	AB025901 Aspergill

26	35	74.5	426	8	ADH22360	Adh22360 Human rec
27	35	74.5	436	8	ADH34619	Adh34619 Endotheli
28	35	74.5	441	2	AAR25394	Aar25394 Sequence
29	35	74.5	442	2	AAR30886	Aar30886 Etb recep
30	35	74.5	442	4	ABB56352	Abb56352 Non-endog
31	35	74.5	442	5	ABR57085	Abbr57085 Mouse isc
32	35	74.5	442	6	ABR58526	Abbr58526 Human end
33	35	74.5	442	6	ABG74670	Abg74670 Human EDN
34	35	74.5	442	6	ABP81815	Abp81815 Human end
35	35	74.5	442	6	ABU61640	Abu61640 Human end
36	35	74.5	442	7	ADE61732	Ad61732 Human pro
37	35	74.5	442	7	ADE61730	Ad61730 Rat Prote
38	35	74.5	442	7	ADE61736	Ad61736 Human pro
39	35	74.5	442	7	ADD46389	Ad46389 Rat Prote
40	35	74.5	442	7	ADD46391	Ad46391 Human pro
41	35	74.5	442	7	AD029303	Ad029303 Human GPC
42	35	74.5	442	7	AD029304	Ad029304 Mouse GPC
43	35	74.5	442	8	ADQ18283	Adq18283 Human sof
44	35	74.5	442	8	ADQ18283	Adq18283 Human sof
45	35	74.5	442	8	ADQ18283	Adq18283 Human sof

#### ALIGNMENTS

RESULT 1	AA022983	standard; peptide; 10 AA.
ID	AA022983	
AC	AA022983	
XX		
DT	17-SEP-2003	(first entry)
DE	Human p280-89 ALK-derived HLA-A*0201 restricted CTL epitope peptide.	
HLA-A*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL; cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human; oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy; t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma; p280-89.		
XX		
OS	Homo sapiens.	
XX		
PN	W02003042243-A2.	
XX		
PD	22-MAY-2003.	
XX		
EF	14-NOV-2002; 2002MO-EP012764.	
XX		
PR	15-NOV-2001; 2001US-00008377.	
XX		
PA	(NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.	
XX	(STUD-) 1ST JUDE CHILDREN'S RES HOSPITAL.	
PI	Gambacorti-Passerini C, Passoni L;	
XX	WPI; 2003-441791/41.	
DR		
PT	New HLA-A*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,	
PT	useful for preparing a composition for treating ALK-positive lymphoma,	
PT	neuroblastoma or ALK-expressing neoplasia.	
XX		
PS	Claim 1; Page 3; 33pp; English.	
XX		
CC	The invention relates to a novel HLA-A*0201-binding anaplastic lymphoma	
CC	kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a	
CC	cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase	
CC	which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase	
CC	fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell	
CC	lines. More than 50% of ALCL cases possess a t(2;5) chromosomal	
CC	translocation that leads to the expression of the NPM/ALK fusion protein	
CC	which forms a potent oncogene when constitutively activated. Translocated	
CC	ALK is a widely expressed tumour-associated antigen characteristic of ALK	

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytosstatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p280-89  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 XX  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 47; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLAMDILHV 10  
 Db 1 SLAMDILHV 10

## RESULT 2

AA873862  
 ID AAB73862 standard; peptide; 45 AA.

AC AAB73862;

DT 15-MAY-2001 (first entry)

DE Human ALK catalytic domain sequence #3.

XX Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;

KW Chromosomal translocation; cancer; NPM/ALK; catalytic domain.

XX Homo sapiens.

OS US6174674-B1.

PI 16-JAN-2001.

PD 19-JUN-1998; 98US-00100089.

XX 03-DEC-1993; 93US-00160861.

PR 12-OCT-1995; 95US-00542363.

XX (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PA Morris SW, Look AT;

XX WPI; 2001-243208/25.

DR WPI; 2001-243208/25.

XX Detection of chromosomal rearrangement or translocations present in t(2;

PT 5) for cancer diagnosis comprises using and identifying anaplastic

XX lymphoma kinase and nucleophosmin genes in nucleic acid hybridization

XX techniques.

XX Example 2; Fig 2C; 87pp; English.

PS The present sequence is given in a specification relating to a method for

XX detecting a chromosomal rearrangement involving a breakpoint in the

CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method

CC is based on the identification and sequence determination of a novel

CC gene, ALK, which is fused to the gene encoding NPM in translocations

CC present in t(2;5) lymphoma cells. The method is useful for detecting a

CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement

CC involving the ALK gene or NPM gene. It is particularly useful in

CC determining if particular cells or tissues express ALK or NPM/ALK coding

CC sequences. The method is also useful in diagnostic assays to determine,

CC for example, if a mammal has cancer or a genetic predisposition to cancer

XX Sequence 45 AA;

SQ Query Match 100.0%; Score 47; DB 4; Length 45;

Best Local Similarity 100.0%; Pred. No. 0.069;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLAMDILHV 10  
 Db 5 SLAMDILHV 14

## RESULT 3

ABG74476  
 ID ABG74476 standard; peptide; 45 AA.

AC ABG74476;

DT 11-APR-2003 (first entry)

DE Tyrosine kinase insulin receptor ALK fragment SEQ ID 14.

XX ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;

KW t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;

XX nucleolar phosphoprotein; centromeric; telomeric.

XX Homo sapiens.

OS US6451997-B1.

PI 17-SEP-2002.

PD 28-SEP-2000; 2000US-00670827.

XX 03-DEC-1993; 93US-00160861.

PR 12-OCT-1995; 95US-00542363.

PR 19-JUN-1998; 98US-00100089.

XX (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PA Morris SW, Look AT;

XX WPI; 2003-101739/09.

DR WPI; 2003-101739/09.

XX Kit for use in method of detecting t(2;5) chromosomal rearrangements or

PT rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic

XX lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK

XX genes.

XX Example 1; Fig 2C; 87pp; English.

PS This invention describes a novel kit for use in a method of detecting t(2;

CC ;5) chromosomal rearrangements (CR) or CRs involving nucleolar

CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method

CC comprises labelled probes that hybridize to a sequence of 2442 or 6226

CC nucleotides or their complement, and to region of human chromosome 5/2

CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to

CC nucleolus). This sequence represents a polypeptide sequence described in

CC the disclosure of the invention

XX Sequence 45 AA;

SQ Query Match 100.0%; Score 47; DB 6; Length 45;

Best Local Similarity 100.0%; Pred. No. 0.069;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDILHV 10

Db 5 SLAMDILHV 14

RESULT 4

AA878521

ID AAR78521 standard; protein; 525 AA.

AC AAR78521;

XX

DT 25-MAR-2003 (revised)  
DT 15-NOV-1995 (first entry)  
XX NPM/ALK fusion protein.  
DE  
XX  
XX ALK protein; protein-tyrosine-kinase; fusion protein; NPM protein;  
KW nuclear phosphoprotein; anaplastic large cell lymphoma; t(2;  
XX 5) translocation; diagnosis; gene therapy.  
XX  
XX Homo sapiens.  
XX  
XX WO9515331-A1.  
XX  
XX 08-JUN-1995.  
XX  
XX 05-DEC-1994; 94WO-US013947.  
XX  
XX 03-DEC-1993; 93US-00160861.  
XX  
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX  
XX Morris SW, Look AT;  
XX  
XX WPI, 1995-215226/28.  
XX DR N-PSDB; AAQ95542.  
XX  
XX Methods for detecting human t(2;5) lymphoma - for detection and diagnosis  
PT of anaplastic large cell lymphoma(s).  
XX  
XX Claim 16; Page 42-44; 70pp; English.  
XX  
XX The translocation event that occurs in human t(2;5) lymphoma brings  
CC sequences from the nuclear phosphoprotein (NPM) gene on chromosome 5q35  
CC to those from a protein-tyrosine-kinase (ALK) gene on chromosome 2q23.  
CC Nucleic acids encoding a partial ALK protein (AAR78520) and the ALK/NPM  
CC fusion protein (AAR78521) were isolated (AAQ95541, AAQ95542).  
CC Identification of the NPM/ALK fusion allows lymphoma diagnosis, or  
CC therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR  
CC 2003 to correct PN field.)  
XX  
XX Sequence 525 AA;  
SQ  
Query Match 100.0%; Score 47; DB 2; Length 525;  
Best Local Similarity 100.0%; Pred. No. 0.99;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 1 SLAMDLHLHV 10  
DB 280 SLAMDLHLHV 289  
RESULT 5  
ABU04368  
ID ABU04368 standard; protein; 525 AA.  
XX  
XX ABU04368;  
XX  
XX 29-JAN-2003 (first entry)  
XX  
XX Human expressed protein tag (EPT) #1034.  
XX  
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
XX Homo sapiens.  
XX  
XX WO200278524-A2.  
XX  
XX 10-OCT-2002.  
XX

EF 28-MAR-2002; 2002MO-US009671.  
XX  
XX 28-MAR-2001; 2001US-0279495P.  
XX  
XX 21-MAY-2001; 2001US-0292544P.  
XX  
XX 08-AUG-2001; 2001US-0310801P.  
XX  
XX 01-OCT-2001; 2001US-0326370P.  
XX  
XX 04-DEC-2001; 2001US-0336780P.  
XX  
XX 20-FEB-2002; 2002US-0358985P.  
XX  
XX (ZYCO-) ZYCO INC.  
XX  
XX Chicx RM, Tomlinson AJ, Urban RG;  
XX  
XX WPI, 2003-040607/03.  
XX  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
XX Example 2; SEQ ID NO 1034; 134pp; English.  
XX  
XX The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 525 AA;  
SQ  
Query Match 100.0%; Score 47; DB 6; Length 525;  
Best Local Similarity 100.0%; Pred. No. 0.99;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 1 SLAMDLHLHV 10  
DB 280 SLAMDLHLHV 289  
RESULT 6  
AAR78520  
ID AAR78520 standard; protein; 572 AA.  
XX  
XX AAR78520;  
XX  
XX 25-MAR-2003 (revised)  
XX  
XX 15-NOV-1995 (first entry)  
XX  
XX Partial ALK protein.  
XX  
XX ALK protein; protein-tyrosine-kinase; fusion protein; NPM protein;  
KW nuclear phosphoprotein; anaplastic large cell lymphoma; t(2;  
KW 5) translocation; diagnosis; gene therapy.  
XX  
XX Homo sapiens.  
XX  
XX WO9515331-A1.  
XX  
XX 08-JUN-1995.  
XX  
XX 05-DEC-1994; 94WO-US013947.  
XX

```

XX 03-DEC-1993; 93US-00160861.
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Morris SW, Look AT;
XX
XX WPI; 1995-215226/28.
XX N-PSDB; AAQ95541.
XX
XX Methods for detecting human t(2;5) lymphoma - for detection and diagnosis
XX of anaplastic large cell lymphoma(s).
XX
XX Claim 21; Page 40-42; 70pp; English.
XX
XX The translocation event that occurs in human t(2;5) lymphoma brings
XX sequences from the nucleolar phosphoprotein (NPM) gene on chromosome 5q35
XX to those from a protein-tyrosine-kinase (ALK) gene on chromosome-2q23.
XX Nucleic acids encoding a partial ALK protein (AA878520) and the ALK/NPM
XX fusion protein (AA878521) were isolated (AAQ95541, AAQ95542).
XX Identification of the NPM/ALK fusion allows lymphoma diagnosis, or
XX therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR
XX -2003 to correct PN field.)
XX
XX Sequence 572 AA;
XX
XX Query Match 100.0%; Score 47; DB 2; Length 572;
XX Best Local Similarity 100.0%; Pred. No. 1.1;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SLAMDLHLV 10
XX 327 SLAMDLHLV 336
XX
XX RESULT 7
XX ABG20950
XX ID ABG20950 standard; protein; 675 AA.
XX
XX AC ABG20950;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #20941.
XX
XX KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US008631.
XX
XX PR 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS85137.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 51309; 103pp; English.

```

```

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 675 AA;
XX
XX Query Match 100.0%; Score 47; DB 4; Length 675;
XX Best Local Similarity 100.0%; Pred. No. 1.3;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SLAMDLHLV 10
XX 275 SLAMDLHLV 284
XX
XX RESULT 8
XX AAB73858
XX ID AAB73858 standard; protein; 680 AA.
XX
XX AC AAB73858;
XX
XX DT 15-MAY-2001 (first entry)
XX
XX DE Human NPM/ALK fusion protein.
XX
XX KM Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;
XX chromosomal translocation; cancer; NPM/ALK; fusion.
XX
XX OS Homo sapiens.
XX
XX PN US6174674-B1.
XX
XX PD 16-JAN-2001.
XX
XX PF 19-JUN-1998; 98US-00100089.
XX
XX PR 03-DEC-1993; 93US-00160861.
XX 12-OCT-1995; 95US-00542363.
XX
XX PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX PI Morris SW, Look AT;
XX
XX WPI; 2001-243208/25.
XX N-PSDB; AAF76868.
XX
XX Detection of chromosomal rearrangement or translocations present in t(2;
XX 5) for cancer diagnosis comprises using and identifying anaplastic
XX lymphoma kinase and nucleophosmin genes in nucleic acid hybridization
XX techniques.
XX
XX Claim 1; Fig 2A; 87pp; English.
XX
XX The present sequence is given in a specification relating to a method for
XX detecting a chromosomal rearrangement involving a breakpoint in the

```

CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 XX

Sequence 680 AA;

Query Match 100.0%; Score 47; DB 4; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDDLHV 10  
 Db 280 SLAMDDLHV 289

RESULT 9  
 ABG95067  
 ID ABG95067 standard; protein; 680 AA.  
 XX  
 AC ABG95067;  
 XX

DT 04-DEC-2002 (first entry)  
 XX  
 DE Human translocation (2; 5) (p23; q35) protein.  
 XX

XX Chromosome aberration; oncogenic fusion protein; cancer;  
 XX Proliferative disease; cellular protein isoform; heat shock protein 90;  
 KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;  
 KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;  
 KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;  
 KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;  
 KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;  
 KW rhabdomyosarcoma; synovial sarcoma; viral infection.  
 XX

OS Homo sapiens.  
 XX  
 XX WO200269900-A2.  
 XX  
 PD 12-SEP-2002.  
 XX

XX 01-MAR-2002; 2002WO-US006518.  
 XX  
 PF 01-MAR-2001; 2001US-0272751P.  
 XX

PR (CONF-) CONFORMA THERAPEUTICS CORP.  
 XX  
 PA Fritz LC, Burrows RJ;  
 XX

PI WPI: 2002-698710/75.  
 XX  
 DR N-PSDB; ABS73246.  
 XX

PT Treating genetically-defined disease associated with chromosomal  
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative  
 PT diseases, involves administering an inhibitor of heat shock protein 90.  
 XX

PS Disclosure; Page 200-202; 369pp; English.

CC The invention describes a method of treating genetically-defined disease  
 CC associated with chromosomal aberrations yielding oncogenic fusion  
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous  
 CC cell population, treating proliferative diseases associated with mutant  
 CC protein or cellular protein isoforms (II) dependent on heat shock protein  
 CC (HSP)-90, or selectively treating cells expressing (II) involving  
 CC administering HSP90-inhibitor. The method is useful for treating  
 CC genetically-defined disease with chromosomal aberration yielding  
 CC oncogenic fusion protein, treating cancerous cells containing fusion  
 CC protein in heterogeneous cell population, treating proliferative disease

CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or  
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.  
 CC P53), or selectively treating cells expressing mutant protein or cellular  
 CC protein isoform in a patient heterozygous for (II). The method is useful  
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell  
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML,  
 CC or a disease characterised by a solid tumour such as papillary thyroid  
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and  
 CC synovial sarcoma. The method is also useful for treating viral  
 CC infections. This represents a protein encoded by the DNA sequence of a  
 CC chromosome aberration  
 XX

Query Match 100.0%; Score 47; DB 5; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDDLHV 10  
 Db 280 SLAMDDLHV 289

RESULT 10  
 ABU04347  
 ID ABU04347 standard; protein; 680 AA.  
 XX  
 AC ABU04347;  
 XX

DT 29-JAN-2003 (first entry)  
 XX

DE Human expressed protein tag (EPT) #1013.  
 XX

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX

OS Homo sapiens.  
 XX  
 XX WO200278524-A2.  
 XX  
 PD 10-OCT-2002.  
 XX

XX 28-MAR-2002; 2002WO-US009671.  
 XX  
 PF 28-MAR-2001; 2001US-0279495P.  
 XX

PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX

PA (ZYCO-) ZYCOS INC.  
 XX

PI Chicx RM, Tomlinson AJ, Urban RG;  
 XX  
 DR WPI: 2003-040607/03.  
 XX

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX

PS Example 2; SEQ ID NO 1013; 134pp; English.

CC The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this

CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 680 AA;

Query Match 100.0%; Score 47; DB 6; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDDLHV 10  
DB 280 SLAMDDLHV 289

## RESULT 11

ABU04362  
ID ABU04362 standard; protein; 680 AA.

XX AC ABU04362;

XX DT 29-JAN-2003 (first entry)

XX DE Human expressed protein tag (EPT) #1028.

XX KM Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX KM protease; protease inhibitor; transporter; cytoskeletal protein;  
XX KM receptor; transcription factor; cancer; MHC;  
XX KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX OS Homo sapiens.

XX PN WO200278524-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US009671.

XX PR 28-MAR-2001; 2001US-0279495P.

XX PR 21-MAY-2001; 2001US-0292544P.

XX PR 08-AUG-2001; 2001US-0310801P.

XX PR 01-OCT-2001; 2001US-0326370P.

XX PR 04-DEC-2001; 2001US-0336780P.

XX PR 20-FEB-2002; 2002US-0358985P.

XX PA (ZYCO-) ZYCOS INC.

XX PI Chicx RM, Tomlinson AJ, Urban RG;

XX DR WPI; 2003-040607/03.

XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX PT cytoskeletal proteins, receptors or transcription factors), useful for  
XX PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX PT leukemia.

XX PS Example 2; SEQ ID NO 1028; 134bp; English.  
XX CC The invention describes a purified polypeptide, which comprises a  
XX CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX CC transporter, cytoskeletal protein, receptor or transcription factor. The  
XX CC polypeptide is useful as an immunogenic composition for eliciting in a  
XX CC mammal an immunogenic response directed against any of the purified  
XX CC polypeptide. The purified polypeptide, or the antibody that binds to this

CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 680 AA;

Query Match 100.0%; Score 47; DB 6; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDDLHV 10  
DB 280 SLAMDDLHV 289

## RESULT 12

ABU04358  
ID ABU04358 standard; protein; 680 AA.

XX AC ABU04358;

XX DT 29-JAN-2003 (first entry)

XX DE Human expressed protein tag (EPT) #1024.

XX KM Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX KM protease; protease inhibitor; transporter; cytoskeletal protein;  
XX KM receptor; transcription factor; cancer; MHC;  
XX KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX OS Homo sapiens.

XX PN WO200278524-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US009671.

XX PR 28-MAR-2001; 2001US-0279495P.

XX PR 21-MAY-2001; 2001US-0292544P.

XX PR 08-AUG-2001; 2001US-0310801P.

XX PR 01-OCT-2001; 2001US-0326370P.

XX PR 04-DEC-2001; 2001US-0336780P.

XX PR 20-FEB-2002; 2002US-0358985P.

XX PA (ZYCO-) ZYCOS INC.

XX PI Chicx RM, Tomlinson AJ, Urban RG;

XX DR WPI; 2003-040607/03.

XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX PT cytoskeletal proteins, receptors or transcription factors), useful for  
XX PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX PT leukemia.

XX PS Example 2; SEQ ID NO 1024; 134bp; English.  
XX CC The invention describes a purified polypeptide, which comprises a  
XX CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX CC transporter, cytoskeletal protein, receptor or transcription factor. The  
XX CC polypeptide is useful as an immunogenic composition for eliciting in a  
XX CC mammal an immunogenic response directed against any of the purified  
XX CC polypeptide. The purified polypeptide, or the antibody that binds to this



CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

SO Sequence 680 AA;

Query Match 100.0%; Score 47; DB 6; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
 |||||  
 Db 280 SLAMDLHLV 289

#### RESULT 13

ID ABG74472 standard; protein; 680 AA.

AC ABG74472;

DT 11-APR-2003 (first entry)

DE NPM/ALK fusion construct SEQ ID 4.

KM ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;  
 KW t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;  
 KM nucleolar phosphoprotein; centromeric; telomeric.

OS Synthetic.

Key Location/Qualifiers  
 1..117  
 FT Protein /note="NPM-derived protein fragment"  
 104..115  
 FT Domain /note="potential metal binding domain"  
 118..680  
 FT Protein /note="ALK-derived protein fragment"  
 182..437  
 FT Domain /label=ALK\_catalytic\_domain

PN US6451997-B1.

PD 17-SEP-2002.

PF 28-SEP-2000; 2000US-00670827.

PR 03-DEC-1993; 93US-00160861.

PR 12-OCT-1995; 95US-00542363.

PR 19-JUN-1998; 98US-00100089.

PA (SUJD-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Morris SW, Look AT;

DR WPI; 2003-101739/09.

DR N-PSDB; ABG77107.

CC Kit for use in method of detecting t(2;5) chromosomal rearrangements or  
 CC rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic  
 CC lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK  
 CC genes.

Example 1; Fig 2A; 87bp; English.

CC This invention describes a novel kit for use in a method of detecting t(2;  
 CC 5) chromosomal rearrangements (CR) or CRs involving nucleolar  
 CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method  
 CC comprises labelled probes that hybridize to a sequence of 2442 or 6226  
 CC nucleotides or their complement, and to region of human chromosome 5/2  
 CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to  
 CC nucleotide). This sequence represents a polypeptide sequence described in  
 CC the disclosure of the invention

SO Sequence 680 AA;

Query Match 100.0%; Score 47; DB 6; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
 |||||  
 Db 280 SLAMDLHLV 289

#### RESULT 14

ID AAB73857 standard; protein; 1620 AA.

AC AAB73857;

DT 15-MAY-2001 (first entry)

DE Human ALK protein.

KM Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;  
 KW chromosomal translocation; cancer.

OS Homo sapiens.

PN US6174674-B1.

PD 16-JAN-2001.

PF 19-JUN-1998; 98US-00100089.

PR 03-DEC-1993; 93US-00160861.

PR 12-OCT-1995; 95US-00542363.

PA (SUJD-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Morris SW, Look AT;

DR WPI; 2001-243208/25.

DR N-PSDB; AAF76867.

CC Detection of chromosomal rearrangement or translocations present in t(2;  
 CC 5) for cancer diagnosis comprises using and identifying anaplastic  
 CC lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 CC techniques.

Claim 10; Fig 3B; 87bp; English.

CC The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer

SO Sequence 1620 AA;

Query Match 100.0%; Score 47; DB 4; Length 1620;

Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
1220 SLAMDLHLV 1229

RESULT 15  
AA018498  
ID AA018498 standard; protein; 1620 AA.

XX AA018498;

XX 11-OCT-2002 (first entry)

DE Human insulin receptor signaling modifier SEQ ID NO: 18.

XX Human; insulin receptor signaling; insulin receptor signaling modifier;

KW ISM; diabetes; metabolic syndrome; antidiabetic.

XX Homo sapiens.

XX WO200255664-A2.

XX 18-JUL-2002.

XX 11-JAN-2002; 2002WO-US001048.

XX 12-JAN-2001; 2001US-0261226P.

XX 12-JAN-2001; 2001US-0261303P.

XX 12-JAN-2001; 2001US-0261304P.

XX 12-JAN-2001; 2001US-0261335P.

XX 12-JAN-2001; 2001US-0261351P.

XX 12-JAN-2001; 2001US-0261456P.

XX 12-JAN-2001; 2001US-0261457P.

XX 12-JAN-2001; 2001US-0261458P.

XX 12-JAN-2001; 2001US-0261459P.

XX 12-JAN-2001; 2001US-0261461P.

XX 12-JAN-2001; 2001US-0261518P.

XX 12-JAN-2001; 2001US-0261531P.

XX 12-JAN-2001; 2001US-0261532P.

XX 12-JAN-2001; 2001US-0261589P.

XX 12-JAN-2001; 2001US-0261590P.

XX 12-JAN-2001; 2001US-0261634P.

XX 12-JAN-2001; 2001US-0261635P.

XX 12-JAN-2001; 2001US-0261697P.

XX (EXEL-) EXELIXIS INC.

XX Seidel-Dugan C, Ferguson KC, Kidd T;

XX WPI, 2002-599664/64.

XX N-PSDB; AAL48617.

PT Identifying an insulin receptor signaling modulator, useful as drug  
PT targets for treating diabetes or metabolic disorders, comprises,  
PT contacting an assay system comprising insulin receptor signaling  
PT modifiers with a test agent.

XX Disclosure; Page 59-66; 232pp; English.

XX The present invention relates to a method of identifying a candidate  
XX insulin receptor (INR) signaling modulating agent, involving contacting  
XX an assay system comprising an insulin receptor signaling modifier (ISM)  
XX polypeptide or nucleic acid with a test agent, and detecting a test agent  
XX -biased activity of the assay system. The method is useful for  
XX identifying candidate INR signaling modulating agents. ISM genes may be  
XX used as drug targets for treatment of disorders related to INR signaling  
XX such as diabetes or metabolic syndrome. ISM nucleic acids and  
XX polypeptides are useful for identifying and testing agents that modulate  
XX ISM function and for other applications related to the involvement of ISM

CC in INR signaling, and for identifying subjects having a predisposition to  
CC such diseases associated with INR signaling. The present sequence is an  
CC ISM protein described in the exemplification of the invention

XX Sequence 1620 AA;

QY Query Match 100.0%; Score 47; DB 5; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 SLAMDLHLV 10  
1220 SLAMDLHLV 1229

Search completed: November 30, 2004, 07:39:53  
Job time : 79.4295 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 30, 2004, 07:19:22 ; Search time 376.923 Seconds  
(without alignments)  
9.410 Million cell updates/sec

Title: US-10-008-377A-1  
Perfect score: 47  
Sequence: 1 SLAMDILHV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database : Published Applications-AA:

1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	10	14	US-10-008-377-1
2	47	100.0	45	15	US-09-827-949-14
3	47	100.0	370	16	US-10-654-421-70
4	47	100.0	680	9	US-09-827-949-4
5	47	100.0	1620	9	US-09-827-949-2
6	43	91.5	9	14	US-10-008-377-3
7	39	83.0	9	14	US-10-008-377-4
8	37	78.7	391	14	US-10-369-493-2155
9	36	76.6	49	15	US-10-424-599-237107
10	36	76.6	129	17	US-10-425-115-202166
11	36	76.6	140	16	US-10-767-701-53326
12	36	76.6	253	17	US-10-425-115-297607
13	36	76.6	376	15	US-10-041-018-26

14	36	76.6	376	15	US-10-041-018-41	Sequence 41, Appl
15	36	76.6	376	15	US-10-041-018-67	Sequence 67, Appl
16	36	76.6	376	16	US-10-437-963-176494	Sequence 176494,
17	36	76.6	563	17	US-10-425-115-297605	Sequence 297605,
18	35	74.5	187	9	US-09-778-927A-68	Sequence 68, Appl
19	35	74.5	302	14	US-10-128-714-3559	Sequence 3559, Ap
20	35	74.5	334	14	US-10-128-714-8559	Sequence 8559, Ap
21	35	74.5	442	10	US-09-826-109-497	Sequence 497, App
22	35	74.5	442	14	US-10-020-141-6	Sequence 6, Appl
23	35	74.5	442	14	US-10-282-525-10	Sequence 10, Appl
24	35	74.5	442	14	US-10-285-567A-114	Sequence 114, Appl
25	35	74.5	442	15	US-10-372-683-49	Sequence 49, Appl
26	35	74.5	532	15	US-10-311-671-11	Sequence 11, Appl
27	35	74.5	532	16	US-10-408-765A-2102	Sequence 2102, Appl
28	34	72.3	496	15	US-10-335-977-5681	Sequence 5681, Ap
29	33	70.2	98	16	US-10-767-701-61922	Sequence 61922, A
30	33	70.2	144	15	US-10-424-599-172565	Sequence 172565,
31	33	70.2	152	15	US-10-425-114-45278	Sequence 45278, A
32	33	70.2	175	15	US-10-282-122A-45446	Sequence 45446, A
33	33	70.2	176	15	US-10-282-122A-46491	Sequence 46491, A
34	33	70.2	292	14	US-10-017-161-32	Sequence 32, Appl
35	33	70.2	292	15	US-10-343-650A-572	Sequence 572, App
36	33	70.2	323	9	US-09-886-055-163	Sequence 163, App
37	33	70.2	323	10	US-09-804-291-163	Sequence 163, App
38	33	70.2	323	11	US-09-981-666A-67	Sequence 67, Appl
39	33	70.2	323	14	US-10-387-629-78	Sequence 78, Appl
40	33	70.2	323	14	US-10-282-798-634	Sequence 634, App
41	33	70.2	323	15	US-10-343-650A-314	Sequence 314, App
42	33	70.2	323	17	US-10-149-826-34	Sequence 34, Appl
43	33	70.2	326	15	US-10-311-671-17	Sequence 17, Appl
44	33	70.2	328	14	US-10-010-568-10	Sequence 10, Appl
45	33	70.2	328	14	US-10-010-568-14	Sequence 14, Appl

## ALIGNMENTS

## RESULT 1

US-10-008-377-1  
Sequence 1, Application US/10008377  
Publicat ion No. US20030157101A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
TITLE OF INVENTION: Immunogenic ALK peptides  
FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008.377  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-1

Query Match 100.0%; Score 47; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.046;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 1 SLAMDILHV 10  
DB 1 SLAMDILHV 10  
RESULT 2  
US-09-827-949-14  
Sequence 14, Application US/09827949  
Patent No. US2001021505A1  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof

```
FILE REFERENCE: 0656.0400004
CURRENT APPLICATION NUMBER: US/09/827,949
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: US 09/670,827
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 09/100,089
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: US 08/542,363
PRIOR FILING DATE: 1995-10-12
PRIOR APPLICATION NUMBER: US 08/160,861
PRIOR FILING DATE: 1993-12-03
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.0
SEQ ID NO 14
LENGTH: 45
TYPE: PRT
ORGANISM: Homo sapiens
US-09-827-949-14
```

```
Query Match          100.0%; Score 47; DB 9; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SLAMDLHLHV 10
      |||||
DB      5 SLAMDLHLHV 14
```

```
RESULT 3
US-10-664-421-70
Sequence 70, Application US/10664421
Publication No. US20040142864A1
GENERAL INFORMATION:
APPLICANT: BREMER, RYAN
APPLICANT: IBRAHIM, BRABHA
APPLICANT: KUMAR, ABHINAV
APPLICANT: MANDIVAN, VALSAN V.
APPLICANT: MILBURN, MICHAEL V.
TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE
FILE REFERENCE: 039363/0703
CURRENT APPLICATION NUMBER: US/10/664,421
CURRENT FILING DATE: 2003-09-16
PRIOR APPLICATION NUMBER: 60/412,341
PRIOR FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: 60/411,398
PRIOR FILING DATE: 2002-09-16
NUMBER OF SEQ ID NOS: 169
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 70
LENGTH: 370
TYPE: PRT
ORGANISM: Homo sapiens
US-10-664-421-70
```

```
Query Match          100.0%; Score 47; DB 16; Length 370;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SLAMDLHLHV 10
      |||||
DB      155 SLAMDLHLHV 164
```

```
RESULT 4
US-09-827-949-4
Sequence 4, Application US/09827949
Publication No. US20010021505A1
GENERAL INFORMATION:
APPLICANT: Morris, Stephan W.
APPLICANT: Look, A. Thomas
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof
FILE REFERENCE: 0656.0400004
CURRENT APPLICATION NUMBER: US/09/827,949
```

```
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: US 09/670,827
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 09/100,089
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: US 08/542,363
PRIOR FILING DATE: 1995-10-12
PRIOR APPLICATION NUMBER: US 08/160,861
PRIOR FILING DATE: 1993-12-03
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 680
TYPE: PRT
ORGANISM: Homo sapiens
US-09-827-949-4
```

```
Query Match          100.0%; Score 47; DB 9; Length 680;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SLAMDLHLHV 10
      |||||
DB      280 SLAMDLHLHV 289
```

```
RESULT 5
US-09-827-949-2
Sequence 2, Application US/09827949
Publication No. US20010021505A1
GENERAL INFORMATION:
APPLICANT: Morris, Stephan W.
APPLICANT: Look, A. Thomas
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof
FILE REFERENCE: 0656.0400004
CURRENT APPLICATION NUMBER: US/09/827,949
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: US 09/670,827
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 09/100,089
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: US 08/542,363
PRIOR FILING DATE: 1995-10-12
PRIOR APPLICATION NUMBER: US 08/160,861
PRIOR FILING DATE: 1993-12-03
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 1620
TYPE: PRT
ORGANISM: Homo sapiens
US-09-827-949-2
```

```
Query Match          100.0%; Score 47; DB 9; Length 1620;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SLAMDLHLHV 10
      |||||
DB      1220 SLAMDLHLHV 1229
```

```
RESULT 6
US-10-008-377-3
Sequence 3, Application US/10008377
Publication No. US20030157101A1
GENERAL INFORMATION:
APPLICANT: Gambacorti-Passerini, Carlo
APPLICANT: Passoni, Lorenza
TITLE OF INVENTION: Immunogenic ALK Peptides
FILE REFERENCE: 045922/241203
CURRENT APPLICATION NUMBER: US/10/008,377
CURRENT FILING DATE: 2001-11-15
```

NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-008-377-3

Query Match 91.5%; Score 43; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAMLDLHV 10  
Db 1 LAMLDLHV 9

RESULT 7  
US-10-008-377-4  
; Sequence 4, Application US/10008377  
; Publication No. US20030157101A1  
; GENERAL INFORMATION:  
; APPLICANT: Gambacorti-Passerini, Carlo  
; APPLICANT: Passoni, Lorenza  
; TITLE OF INVENTION: Immunogenic Alk Peptides  
; FILE REFERENCE: 045922/241203  
; CURRENT APPLICATION NUMBER: US/10/008,377  
; CURRENT FILING DATE: 2001-11-15  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-008-377-4

Query Match 83.0%; Score 39; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LAMLDLHV 10  
Db 1 LAMLDLHV 8

RESULT 8  
US-10-369-493-2155  
; Sequence 2155, Application US/10369493  
; Publication No. US2003032367SA1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xiaofeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10152052/B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2155  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Schizosaccharomyces pombe  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)...(391)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-2155

Query Match 78.7%; Score 37; DB 14; Length 391;  
Best Local Similarity 88.9%; Pred. No. 1.3e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LAMLDLHV 10  
Db 125 LAMLDLHV 133

RESULT 9  
US-10-424-599-237107  
; Sequence 237107, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223) B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 237107  
; LENGTH: 49  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_56132C.1.pep  
US-10-424-599-237107

Query Match 76.6%; Score 36; DB 15; Length 49;  
Best Local Similarity 70.0%; Pred. No. 24;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SLAMLDLHV 10  
Db 20 SLAMLDLHV 29

RESULT 10  
US-10-425-115-202166  
; Sequence 202166, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223) B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 202166  
; LENGTH: 129  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_115965C.1.pep  
US-10-425-115-202166

Query Match 76.6%; Score 36; DB 17; Length 129;  
Best Local Similarity 80.0%; Pred. No. 63;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SLAMLDLHV 10  
Db 81 SLAMLDLHV 90

## RESULT 11

US-10-767-701-53329  
; Sequence 53329, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(5353)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 53329  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(140)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 13318357.pcp  
US-10-767-701-53329

## Query Match

Best Local Similarity 76.6%; Score 36; DB 16; Length 140;  
Best Local Similarity 66.7%; Pred. No. 69;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
DB 131 LAMMDLHV 139

## RESULT 12

US-10-425-115-297607  
; Sequence 297607, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 297607  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(253)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_34493C.1.pcp  
US-10-425-115-297607

## Query Match

Best Local Similarity 76.6%; Score 36; DB 17; Length 253;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
DB 186 LAMMDLHV 194

RESULT 13  
US-10-041-018-26

; Sequence 26, Application US/10041018  
; Publication No. US20040072323A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seichi P.T.  
; APPLICANT: Hart, Elizabeth A.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080US1/10025547  
; CURRENT APPLICATION NUMBER: US/10/041,018  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259880  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-041-018-26

## Query Match

Best Local Similarity 76.6%; Score 36; DB 15; Length 376;  
Best Local Similarity 70.0%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLHV 10  
DB 250 SLAMLEYHV 259

## RESULT 14

US-10-041-018-41  
; Sequence 41, Application US/10041018  
; Publication No. US20040072323A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seichi P.T.  
; APPLICANT: Hart, Elizabeth A.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080US1/10025547  
; CURRENT APPLICATION NUMBER: US/10/041,018  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259880  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 41  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-041-018-41

## Query Match

Best Local Similarity 76.6%; Score 36; DB 15; Length 376;  
Best Local Similarity 70.0%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLHV 10  
DB 250 SLAMLEYHV 259

## RESULT 15

US-10-041-018-67  
; Sequence 67, Application US/10041018  
; Publication No. US20040072323A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seichi P.T.  
; APPLICANT: Hart, Elizabeth A.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080US1/10025547  
; CURRENT APPLICATION NUMBER: US/10/041,018  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259880  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: PatentIn version 3.1

Tue Nov 30 08:49:19 2004

us-10-008-377a-1.open.rapb

Page 5

; SEQ ID NO 67  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-041-018-67

Query Match 76.6%; Score 36; DB 15; Length 376;  
Best Local Similarity 70.0%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 1 SLAMDLYHV 10  
|||: ||  
DB 250 SLAMDLYTHV 259

Search completed: November 30, 2004, 08:06:34  
Job time : 377.923 secs

**This Page Blank (uspio)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12, Search time 25.182 Seconds

(without alignments)  
26.392 Million cell updates/sec

Title: US-10-008-377A-1

Sequence: 1 SLAMDLHLV 10

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:  
1: /cgn2\_6/prodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/ECTUS.COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	100.0	45	1	US-08-542-363-14
2	47	100.0	45	3	US-09-100-089-14
3	47	100.0	45	4	US-09-670-827-14
4	47	100.0	45	4	US-09-827-949-14
5	47	100.0	525	1	US-08-160-861-4
6	47	100.0	572	1	US-08-160-861-3
7	47	100.0	680	1	US-08-542-363-4
8	47	100.0	680	3	US-09-100-089-4
9	47	100.0	680	4	US-09-670-827-4
10	47	100.0	680	4	US-09-827-949-4
11	47	100.0	1620	1	US-08-542-363-2
12	47	100.0	1620	3	US-09-100-089-2
13	47	100.0	1620	4	US-09-670-827-2
14	47	100.0	1620	4	US-09-827-949-2
15	47	100.0	442	3	US-08-121-446-4
16	47	100.0	442	4	US-09-520-210-10
17	47	100.0	442	1	US-08-117-361C-1
18	47	100.0	145	4	US-09-517-999C-7009
19	47	100.0	153	4	US-09-252-991A-21109
20	47	100.0	333	4	US-08-118-270-40
21	47	100.0	333	5	PCR-US93-08528-40
22	47	100.0	294	4	US-09-543-681A-4399
23	47	100.0	385	4	US-09-489-039A-11317
24	47	100.0	387	4	US-09-543-681A-7130
25	47	100.0	391	4	US-09-270-767-43901
26	47	100.0	404	4	US-09-252-991A-18789
27	47	100.0	513	4	US-09-252-991A-30773

28	32	68.1	1047	4	US-09-529-239D-99	Sequence 99, Appl
29	31	66.0	236	1	US-08-015-985-6	Sequence 6, Appl
30	31	66.0	236	4	US-09-280-597-6	Sequence 6, Appl
31	31	66.0	254	2	US-08-685-992-9	Sequence 9, Appl
32	31	66.0	254	2	US-09-144-925-9	Sequence 9, Appl
33	31	66.0	340	4	US-09-712-368-1	Sequence 1, Appl
34	31	66.0	387	3	US-09-134-001C-3625	Sequence 3625, Ap
35	31	66.0	401	4	US-09-252-991A-19899	Sequence 19899, A
36	31	66.0	476	4	US-09-543-681A-6621	Sequence 6621, Ap
37	31	66.0	576	4	US-09-489-039A-7859	Sequence 7859, Ap
38	31	66.0	781	4	US-09-252-991A-26207	Sequence 26207, A
39	31	66.0	793	1	US-08-015-985-3	Sequence 3, Appl
40	31	66.0	793	4	US-09-280-597-3	Sequence 3, Appl
41	31	66.0	802	1	US-08-015-985-1	Sequence 1, Appl
42	31	66.0	802	4	US-09-280-597-1	Sequence 1, Appl
43	31	66.0	1032	4	US-09-252-991A-24058	Sequence 24058, A
44	31	66.0	1101	3	US-09-770-170-8	Sequence 8, Appl
45	31	66.0	2713	5	PCT-US96-01735-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-08-542-363-14  
Sequence 14, Application US/08542363  
Patent No. 5770421

GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.

APPLICANT: Look, A. Thomas

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and

TITLE OF INVENTION: Ligands Thereof

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESSES:

ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/542,363

FILING DATE: 12-OCT-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400001/SLF/GKT

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 45 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-542-363-14

Query Match

Best Local Similarity 100.0%, Pred. No. 0.096, Length 45;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10

DB 5 SLAMDLHLV 14

RESULT 2  
US-09-100-089-14  
; Sequence 14, Application US/09100089  
; Patent No. 6174674  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Stephan W.  
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
; TITLE OF INVENTION: Ligands Thereof  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/100,089  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/160,861  
; FILING DATE: 03-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/542,363  
; FILING DATE: 12-OCT-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fox, Samuel L.  
; REGISTRATION NUMBER: 30,353  
; REFERENCE/DOCKET NUMBER: 0656,0400002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-100-089-14  
Query Match 100.0%; Score 47; DB 3; Length 45;  
Best Local Similarity 100.0%; Pred. No. 0.096;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLAMDLHLV 10  
DB 5 SLAMDLHLV 14  
RESULT 3  
US-09-670-827-14  
; Sequence 14, Application US/09670827  
; Patent No. 6451997  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Stephan W.  
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
; TITLE OF INVENTION: Ligands Thereof  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: DC

COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-Sep-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
APPLICATION NUMBER: US 09/100,089  
FILING DATE: 19-JUN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656,0400003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-670-827-14  
Query Match 100.0%; Score 47; DB 4; Length 45;  
Best Local Similarity 100.0%; Pred. No. 0.096;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLAMDLHLV 10  
DB 5 SLAMDLHLV 14  
RESULT 4  
US-09-827-949-14  
; Sequence 14, Application US/09827949  
; Patent No. 6696548  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Stephan W.  
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
; FILE REFERENCE: 0656,0400004  
; CURRENT APPLICATION NUMBER: US/09/827,949  
; CURRENT FILING DATE: 2001-04-09  
; PRIOR APPLICATION NUMBER: US 09/670,827  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 09/100,089  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: US 08/542,363  
; PRIOR FILING DATE: 1995-10-12  
; PRIOR APPLICATION NUMBER: US 08/160,861  
; PRIOR FILING DATE: 1993-12-03  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 14  
; LENGTH: 45  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-827-949-14  
Query Match 100.0%; Score 47; DB 4; Length 45;  
Best Local Similarity 100.0%; Pred. No. 0.096;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
|||||

Db 5 SLAMDLHLV 14

RESULT 5  
US-08-160-861-4  
Sequence 4, Application US/08160861  
Patent No. 5529295  
GENERAL INFORMATION:  
APPLICANT: MORRIS, STEPHAN W  
TITLE OF INVENTION: NOVEL FUSION NUCLEIC ACID SEQUENCES AND  
TITLE OF INVENTION: FUSION PROTEINS PRESENT IN HUMAN t(2:5) LYMPHOMA, METHODS  
TITLE OF INVENTION: OF DETECTION AND USES THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX  
STREET: 1100 NEW YORK AVE NW SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/160,861  
FILING DATE: 02-DEC-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MILLMAN, ROBERT A  
REGISTRATION NUMBER: 36217  
REFERENCE/DOCKET NUMBER: 0656.0400000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2678  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 525 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-160-861-4

Query Match 100.0%; Score 47; DB 1; Length 525;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
|||||

Db 280 SLAMDLHLV 289

RESULT 6  
US-08-160-861-3  
Sequence 3, Application US/08160861  
Patent No. 5529295  
GENERAL INFORMATION:  
APPLICANT: MORRIS, STEPHAN W  
TITLE OF INVENTION: NOVEL FUSION NUCLEIC ACID SEQUENCES AND  
TITLE OF INVENTION: FUSION PROTEINS PRESENT IN HUMAN t(2:5) LYMPHOMA, METHODS  
TITLE OF INVENTION: OF DETECTION AND USES THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX  
STREET: 1100 NEW YORK AVE NW SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/160,861  
FILING DATE: 02-DEC-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MILLMAN, ROBERT A  
REGISTRATION NUMBER: 36217  
REFERENCE/DOCKET NUMBER: 0656.0400000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2678  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 572 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-160-861-3

Query Match 100.0%; Score 47; DB 1; Length 572;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
|||||

Db 327 SLAMDLHLV 336

RESULT 7  
US-08-542-363-4  
Sequence 4, Application US/08542363  
Patent No. 570421  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,363  
FILING DATE: 12-OCT-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400001/SLF/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 680 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-542-363-4

Query Match 100.0%; Score 47; DB 1; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLAMDLHLV 10  
|||  
Db 280 SLAMDLHLV 289

RESULT 8  
US-09-100-089-4  
; Sequence 4, Application US/09100089  
; Patent No. 6174674  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Stephan W.  
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
; TITLE OF INVENTION: Ligands Thereof  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/100,089  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/160,861  
; FILING DATE: 03-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/542,363  
; FILING DATE: 12-OCT-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fox, Samuel L.  
; REGISTRATION NUMBER: 30,353  
; REFERENCE/DOCKET NUMBER: 0656,0400002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2500  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 680 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-100-089-4

Query Match 100.0%; Score 47; DB 3; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
|||  
Db 280 SLAMDLHLV 289

RESULT 9  
US-09-670-827-4  
; Sequence 4, Application US/09670827  
; Patent No. 6451997  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Stephan W.  
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
; TITLE OF INVENTION: Ligands Thereof  
; NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-Sep-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
APPLICATION NUMBER: US 09/100,089  
FILING DATE: 12-JUN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656,0400003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2500  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 680 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-670-827-4

Query Match 100.0%; Score 47; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
|||  
Db 280 SLAMDLHLV 289

RESULT 10  
US-09-827-949-4  
; Sequence 4, Application US/09827949  
; Patent No. 6696548  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Stephan W.  
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
; FILE REFERENCE: 0656,0400004  
; CURRENT APPLICATION NUMBER: US/09/827,949  
; CURRENT FILING DATE: 2001-04-09  
; PRIOR APPLICATION NUMBER: US 09/670,827  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 09/100,089  
; PRIOR FILING DATE: 1998-05-19  
; PRIOR APPLICATION NUMBER: US 08/542,363  
; PRIOR FILING DATE: 1995-10-12  
; PRIOR APPLICATION NUMBER: US 08/160,861  
; PRIOR FILING DATE: 1993-12-03  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: Patentin Version 3.0  
; SEQ ID NO 4  
; LENGTH: 680  
; TYPE: PRT  
; ORGANISM: Homo sapiens

Tue Nov 30 08:49:18 2004

us-10-008-377a-1.open.rat

Page 5

US-09-827-949-4

Query Match 100.0%; Score 47; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 280 SLAMDLHLV 289

RESULT 11

US-08-542-363-2  
Sequence 2, Application US/08542363  
Patent No. 5770421  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESS: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,363  
FILING DATE: 12-OCT-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400001/SLF/SKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2540  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ. ID NO.: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1620 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-542-363-2

Query Match 100.0%; Score 47; DB 1; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 1220 SLAMDLHLV 1229

RESULT 12  
US-09-100-089-2

Sequence 2, Application US/09100089  
Patent No. 6174674  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ. ID NO.: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1620 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-100-089-2

Query Match 100.0%; Score 47; DB 3; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 1220 SLAMDLHLV 1229

RESULT 13  
US-09-670-827-2

Sequence 2, Application US/09670827  
Patent No. 6431997  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-Sep-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
APPLICATION NUMBER: US 09/100,089  
FILING DATE: 19-JUN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1620 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-670-827-2

Query Match 100.0%; Score 47; DB 4; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 1220 SLAMDLHLV 1229

RESULT 14  
US-09-827-949-2  
Sequence 2, Application US/09827949  
Patent No. 6696548  
GENERAL INFORMATION:  
APPLICANT: Morita, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/670,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO: 2  
LENGTH: 1620  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-2

Query Match 100.0%; Score 47; DB 4; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 1220 SLAMDLHLV 1229

RESULT 15  
US-08-121-446-4  
Sequence 4, Application US/08121446  
Patent No. 6313276  
GENERAL INFORMATION:

APPLICANT: IMURA, HIROO  
APPLICANT: NAKANO, KAZUWA  
APPLICANT: NAKANISHI, SHIGERU  
TITLE OF INVENTION: A HUMAN ENDOTHELIN RECEPTOR  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,446  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/911,684  
FILING DATE: 10-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: CIOTTI, THOMAS E.  
REGISTRATION NUMBER: 21,013  
REFERENCE/DOCKET NUMBER: 29900-20324.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 442 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-121-446-4

Query Match 74.5%; Score 35; DB 3; Length 442;  
Best Local Similarity 70.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 142 SLAMDLHLV 151

Search completed: November 30, 2004, 07:19:15  
Job time : 26.1282 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 20 Seconds

(without alignments)  
48.108 Million cell updates/sec

Title: US-10-008-377A-2

Perfect score: 52

Sequence: 1 GVLMWEIFSL 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1.\*  
2: p1r2.\*  
3: p1r3.\*  
4: p1r4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	367	2 B56598	endothelial kinase
2	52	100.0	464	2 A36625	protein-tyrosine k
3	52	100.0	576	2 S12792	protein-tyrosine k
4	52	100.0	680	4 I38491	nucleophosmin/ana
5	52	100.0	864	2 A48266	protein-tyrosine k
6	52	100.0	888	2 I58378	tyrosine kinase -
7	52	100.0	1088	1 A48993	platelet-derived g
8	52	100.0	1298	2 A48993	protein-tyrosine k
9	52	100.0	1330	2 A49010	embryonic receptor
10	52	100.0	1333	2 I78875	receptor tyrosine
11	52	100.0	1336	2 I60598	Fit-1 tyrosine kin
12	52	100.0	1338	2 S09982	protein-tyrosine k
13	52	100.0	1348	2 S16566	vascular endotheli
14	52	100.0	1356	2 J01402	protein-tyrosine k
15	52	100.0	1363	2 I58375	protein-tyrosine k
16	52	100.0	1367	2 A41228	protein-tyrosine k
17	52	100.0	1379	2 J04954	protein-tyrosine k
18	52	100.0	1621	2 T10200	vascular endotheli
19	51	98.1	1443	2 B39061	protein-tyrosine k
20	51	98.1	160	2 A39061	protein-tyrosine k
21	51	98.1	457	2 S44269	platelet-derived g
22	51	98.1	941	1 TYHWD	protein-tyrosine k
23	51	98.1	954	1 S1703	c-kit-related kina
24	51	98.1	972	1 TYHWD	macrophage colony-
25	51	98.1	975	2 T10816	macrophage colony-
26	51	98.1	976	2 TYHWD	macrophage colony-
27	51	98.1	978	1 S16385	macrophage colony-
28	51	98.1	980	1 TVCTMD	macrophage colony-
29	51	98.1	992	2 A39931	protein-tyrosine k

30	51	98.1	993	2 A36873	protein-tyrosine k
31	51	98.1	1000	2 S18827	Plt3 protein - mou
32	51	98.1	1087	2 I51552	platelet-derived g
33	51	98.1	1089	1 PFHGA	platelet-derived g
34	51	98.1	1089	1 S33727	platelet-derived g
35	50	96.2	478	2 T32476	probable protein-t
36	49	94.2	388	2 I51023	fibroblast growth
37	49	94.2	729	2 A56795	fibroblast growth
38	49	94.2	733	2 I49293	fibroblast growth
39	49	94.2	737	2 S38579	fibroblast growth
40	49	94.2	800	1 TVH2F	fibroblast growth
41	49	94.2	800	2 A48991	heparin-binding gr
42	49	94.2	806	1 TVH0F3	fibroblast growth
43	49	94.2	806	2 A35963	protein-tyrosine k
44	49	94.2	812	1 A36477	fibroblast growth
45	49	94.2	814	1 A39752	fibroblast growth

#### ALIGNMENTS

##### RESULT 1

B56598  
endothelial kinase Quek2 - quail (fragment)  
N.Alternate names: vascular endothelial growth factor receptor homolog Quek2  
C.Species: Coturnix coturnix (quail)  
C.Date: 11-Aug-1995 #sequence\_revision 11-Aug-1995 #ext\_change 16-Aug-2004  
C.Accession: B56598  
R.Reichmann, A.; Kercelle, C.; Breant, C.; Le Douarin, N.M.  
Mech. Dev. 42, 33-48, 1993  
A.Reference: Two molecules related to the VEGF receptor are expressed in early endothelial  
A.Reference number: A56598; MUID:93378866; PMID:8396413  
A.Contents: sep. japonica, E4 embryo  
A.Accession: B56598  
A.Status: Preliminary  
A.Molecule type: mRNA  
A.Residues: 1-367 <SIC>  
A.Cross-references: UNIPROT:Q91356; GB:S65207; NID:G410682; PIDN:AAE28128.1; PID:G41068  
A.Note: sequence extracted from NCBI backbone (NCBIN:137164, NCBI:137165)  
C.Superfamily: protein kinase homology  
C.Keywords: ATP; growth factor receptor  
P.1-176/Domain: protein kinase homology (fragment) <KIN>

Query Match Score 52; DB 2; Length 367;

Best local similarity 100.0%; Pred. No. 0.13;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
|||||||  
DB 101 GVLMWEIFSL 110

##### RESULT 2

A36625  
protein-tyrosine kinase (BC 2.7.1.112) Itk (version 2) - human  
C.Species: Homo sapiens (man)  
C.Date: 11-Feb-1993 #sequence\_revision 11-Feb-1993 #ext\_change 04-Feb-2000  
C.Accession: A43625  
R.Maru, Y.; Hirai, H.; Takaku, F.  
Oncogene Res. 5, 199-204, 1990  
A.Title: Human Itk: gene structure and preferential expression in human leukemic cells.  
A.Reference number: A43625; MUID:90206632; PMID:2320375  
A.Accession: A43625  
A.Status: Preliminary  
A.Molecule type: mRNA  
A.Residues: 1-464 <MAR>  
A.Cross-references: GB:X52213; NID:G34421; PIDN:CAA36460.1; PID:G34422  
C.Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol  
C.Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase  
P.109-384/Domain: protein kinase homology <KIN>  
P.117-125/Region: protein kinase ATP-binding motif

Query Match Score 52; DB 2; Length 464;

Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
DB 310 GVLMWEIFSL 319

## RESULT 3

S12792  
Protein-tyrosine kinase (EC 2.7.1.112) ltk - mouse  
N/Alternate names: leukocyte tyrosine kinase  
C/Species: Mus musculus (house mouse)  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-2004  
C/Accession: S12792; S00904  
R/Bernards, A.; de la Monte, S.M.  
EMBO J. 9, 2279-2287, 1990  
A/Title: The ltk receptor tyrosine kinase is expressed in pre-B lymphocytes and cerebral  
A/Reference number: S12792; PMID:90291994; PMID:2357970  
A/Accession: S12792  
A/Molecule type: mRNA  
A/Residues: 1-576 <DB>  
A/Cross-references: UNIPROT:P08923; EMBL:X52621; NID:G52948; PIDN:CAA36848.1; PID:G43792  
R/Ben-Neriah, Y.; Bauskin, A.R.  
Nature 333, 672-676, 1988  
A/Title: Leukocytes express a novel gene encoding a putative transmembrane protein-kinase  
A/Reference number: S00904; PMID:88232962; PMID:2836739  
A/Accession: S00904  
A/Molecule type: mRNA  
A/Residues: 101-476, 'V', 478-562, 'H', 564-576 <BEN>  
A/Cross-references: EMBL:X07984  
C/Genetics:  
A/Start codon: CTG  
C/Superfamily: protein kinase homology  
C/Keywords: ATP, autophosphorylation, glycoprotein, membrane protein, phosphoprotein, ph  
F/192-468/Domain: protein kinase homology <KIN>  
F/200-208/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 576;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
DB 394 GVLMWEIFSL 403

## RESULT 4

I38491  
nucleophosmin/anaplastic lymphoma kinase mutant fusion protein - human  
C/Species: Homo sapiens (man)  
C/Date: 15-Feb-1996 #sequence\_revision 15-Feb-1996 #text\_change 20-Apr-2000  
C/Accession: I38491  
R/Korits, S.W.; Kiststein, M.N.; Valentine, M.B.; Dittmer, K.G.; Shapiro, D.N.; Salzman,  
Science 263, 1281-1284, 1994  
A/Title: Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin'  
A/Reference number: A53182; PMID:94167588; PMID:8122112  
A/Accession: I38491  
A/Molecule type: mRNA  
A/Residues: 1-560 <NOR>  
A/Cross-references: EMBL:U04946; NID:G609341; PIDN:AAA86848.1; PID:G609342  
C/Comment: This sequence is the chimeric product of a translocation mutation.  
C/Genetics:  
A/Gene: NPM1/ALK  
A/Map position: 5/2p23-2p23  
C/Keywords: fusion protein

Query Match 100.0%; Score 52; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 0.25; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10

DB 376 GVLMWEIFSL 385

## RESULT 5

A48266  
protein-tyrosine kinase (EC 2.7.1.112) ltk - human  
N/Alternate names: protein-tyrosine kinase tyk1  
C/Species: Homo sapiens (man)  
C/Date: 16-Feb-1994 #sequence\_revision 12-Apr-1996 #text\_change 09-Jul-2004  
C/Accession: A48266; S17452; A60189  
R/Toyoshima, H.; Kozutsumi, H.; Maru, Y.; Hagiwara, K.; Futuya, A.; Mich, H.; Hanai, N.;  
Proc. Natl. Acad. Sci. U.S.A. 90, 5404-5408, 1993  
A/Title: Differently spliced cDNAs of human leukocyte tyrosine kinase receptor tyrosine  
A/Reference number: A48266; PMID:93296146; PMID:7685902  
A/Accession: A48266  
A/Molecule type: mRNA  
A/Residues: 1-864 <TOY>  
A/Cross-references: UNIPROT:P29376; GB:D16105; NID:G440854; PIDN:BAAC3679.1; PID:dl00419  
A/Experimental source: placenta  
A/Note: sequence modified after extraction from NCBI backbone  
A/Note: the authors translated the codon CAG for residue 42 as Arg  
A/Note: sequence extracted from NCBI backbone (NCBI:133811)  
R/Krolewski, J.J.; Dalla-Favera, R.  
EMBO J. 10, 2911-2919, 1991  
A/Title: The ltk gene encodes a novel receptor-type protein tyrosine kinase.  
A/Reference number: S17452; PMID:92007725; PMID:1655406  
A/Accession: S17452  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-41, 'R', 43-219, 'L', 221-273, 335-864 <KRO>  
A/Cross-references: EMBL:X60702; NID:G34419; PIDN:CA43113.1; PID:G34420  
R/Krolewski, J.J.; Lee, R.; Eddy, R.; Shows, T.B.; Dalla-Favera, R.  
Oncogene 5, 277-282, 1990  
A/Title: Identification and chromosomal mapping of new human tyrosine kinase genes.  
A/Reference number: A60189; PMID:90191712; PMID:2156206  
A/Accession: A60189  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 608-716 <KR2>  
C/Genetics:  
A/Gene: GDB:LTk  
A/Cross-references: GDB:127768; OMIM:151520  
A/Map position: 15q15.1-15q15.2  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C/Keywords: alternative splicing; ATP; phosphotransferase; tyrosine-specific protein kin  
F/508-784/Domain: protein kinase homology <KIN>  
F/516-524/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 864;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
DB 710 GVLMWEIFSL 719

## RESULT 6

I58378  
tyrosine kinase - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 16-Aug-2004  
C/Accession: I58378  
R/Snijders, A.J.; Haase, V.H.; Bernards, A.  
Oncogene 8, 27-35, 1993  
A/Title: Four tissue-specific mouse ltk mRNAs predict tyrosine kinases that differ upstr  
A/Reference number: I58378; PMID:93141274; PMID:6380920  
A/Accession: I58378  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-888 <RES>  
A/Cross-references: UNIPROT:P08923; GB:M90470; NID:G198885; PIDN:AAA39451.1; PID:G198885



C:Genetics:  
A:Gene: Itk  
C:Superfamily: protein kinase homology  
C:Keywords: ATP  
F:504-780/Domain: protein kinase homology <KIN>  
F:512-520/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 888;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
DB 706 GVLMEIFSL 715

## RESULT 7

PERGTA

platelet-derived growth factor receptor alpha precursor - rat  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text\_change 09-Jul-2004

C:Accession: A34710; S33767; S25100

R:Lee, K.H.; Bowen-Pope, D.P.; Reed, R.R.

Mol. Cell. Biol. 10, 2237-2246, 1990

A>Title: Isolation and characterization of the alpha platelet-derived growth factor rece

A:Reference number: A34710; MUID:90220609; PMID:2157969

A:Accession: A34710

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1088 <LE>

A:CROSS-references: UNIPROT:P20786; GB:M33837; NID:G202929; PIDN:AAA40743.1; PID:G202930

A>Note: in the authors' translation an additional residue, Val, is shown after position

R:Herren, B.; Weyer, K.A.; Rouse, M.; Loetscher, P.; Pech, M.

Biochim. Biophys. Acta 1173, 294-302, 1993

A>Title: Conservation in sequence and affinity of human and rodent PDGF ligands and rece

A:Reference number: S33764; MUID:93305723; PMID:8318539

A:Accession: S33767

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 33-149; 'R', 151-518; 'T', 520-523 <HER1>

A:CROSS-references: EMBL:Z14118; NID:G56863; PIDN:CAA78488.1; PID:G56864

A:Experimental source: strain Sprague Dawley

R:Herren, B.; Weyer, K.A.; Rouse, M.; Loetscher, P.; Pech, M.

submitted to the EMBL Data Library, July 1992

A>Description: Cross-species conservation in sequence and function of PDGF ligands and r

A:Reference number: S25096

A:Accession: S25100

A:Molecule type: mRNA

A:Residues: 33-149; 'R', 151-518; 'T', 520-523 <HER2>

A:CROSS-references: EMBL:Z14118; NID:G56863; PIDN:CAA78488.1; PID:G56864

C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;

C:Keywords: ATP; autophosphorylation; glycoprotein; heterodimer; homodimer; phosphoprote

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-108/Region: product; platelet-derived growth factor receptor alpha #status predicted <MNT>

F:24-523/Domain: extracellular #status predicted <EXT>

F:41-101/Domain: immunoglobulin homology <IMM1>

F:142-190/Domain: immunoglobulin homology <IMM2>

F:227-291/Domain: immunoglobulin homology <IMM3>

F:524-547/Domain: transmembrane #status predicted <TMN4>

F:548-1088/Domain: intracellular #status predicted <INT>

F:598-955/Region: protein kinase homology <KIN>

F:148-991/Region: protein kinase ATP-binding motif

F:175; '6', 88; '102', 178; '35', 457; Binding site: carbohydrate (Asn) (covalent) #status p

F:626/Active site: Lys #status predicted

F:848/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 100.0%; Score 52; DB 1; Length 1088;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
DB 881 GVLMEIFSL 890

## RESULT 8

A48999

protein-tyrosine kinase (EC 2.7.1.112) Flt4 precursor - human

N:Alternate names: class III receptor tyrosine kinase FLT4; fms-like tyrosine kinase 4

C:Species: Homo sapiens (man)

C>Date: 21-Jan-1994 #sequence revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A48999; A44930; G02316; S36130; A42010

R:Paavola, K.; Aprelikova, O.; Korhonen, J.; Kaipainen, A.; Pertovaara, L.; Alitalo, R.

Cancer Res. 52, 5738-5743, 1992

A>Title: FLT4 receptor tyrosine kinase contains seven immunoglobulin-like loops and is

A:Reference number: A48999; MUID:93007958; PMID:1327515

A:Accession: A48999

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1298 <PAJ>

A:CROSS-references: UNIPROT:P35916; PIDN:AA823636.1; PID:G257352

A:Experimental source: HEL erythroleukemia cells

R:Aprelikova, O.; Paavola, K.; Partanen, J.; Armstrong, E.; Alitalo, R.; Bailey, S.K.;

Cancer Res. 52, 746-748, 1992

A>Title: FLT4, a novel class III receptor tyrosine kinase in chromosome 5q31-qter.

A:Reference number: A44930; MUID:92119639; PMID:1310071

A:Accession: A44930

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 761-1190 <APR>

A:CROSS-references: GB:X68203; NID:G31433

A>Note: sequence extracted from NCBI backbone (NCBIP:78155)

R:Wood, W.I.

submitted to the EMBL Data Library, December 1995

A:Reference number: H01039

A:Accession: G02316

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1298 <MOO>

A:CROSS-references: EMBL:U43143; NID:G1150990; PIDN:AAA85215.1; PID:G1150991

R:Galland, F.

submitted to the EMBL Data Library, December 1992

A:Reference number: S36130

A:Accession: S36130

A:Molecule type: mRNA

A:Residues: 1-23; 'D', 25-744; 'P', 746-751; 'RP', 754-869; 'Q', 891-1127; 'V', 1129-1145; 'H', 1147

A:CROSS-references: EMBL:X69878; NID:G257049; PIDN:CAA49505.1; PID:G297050

R:Galland, F.; Karamysheva, A.; Mattei, M.G.; Rosnet, O.; Marchetto, S.; Birnbaum, D.

Genomics 13, 475-478, 1992

A>Title: Chromosomal localization of FLT4, a novel receptor-tyrosine kinase gene.

A:Reference number: A42010; MUID:92307693; PMID:1319394

A:Accession: A42010

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr

A:Molecule type: mRNA

A:Residues: 776-869; 'Q', 891-1127; 'V', 1129-1145; 'H', 1147-1163; 'D', 1165-1200 <GAL2>

C:Genetics:

A:Gene: GDB:FLT4

A:CROSS-references: GDB:128732; OMIM:136352

A:Map position: 5q34-5q35

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; transmembrane

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-1298/Product: protein-tyrosine kinase FLT4 #status predicted <MNT>

F:843-1176/Domain: protein kinase homology <KIN>

F:851-859/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 1298;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
DB 10 GVLMEIFSL 10

Db 1101 GVLMEIFSL 1110

# RESULT 9

549010

embryonic receptor kinase - mouse

C:Species: Mus musculus (house mouse)

C>Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 16-Aug-2004

C:Accession: S49010

R:Choi, K.; Wall, C.; Hanratty, R.; Keller, G.

Oncogene 9, 1261-1266, 1994

A:Title: Isolation of a gene encoding a novel receptor tyrosine kinase from differentiat

A:Reference number: S49010; MUID:94181281; PMID:8134130

A:Accession: S49010

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-1130 <CHO>

A:Cross-references: UNIPROT:P35969; EMBL:X78568; NID:G510664; PIDN:CAAS5311.1; PID:G5106

C:Superfamily: protein kinase homology

C:Keywords: ATP

F:831-839/Region: protein kinase ATP-binding motif

## Query Match

Best Local Similarity 100.0%; Score 52; DB 2; Length 1330;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10

Db 1083 GVLMEIFSL 1092

# RESULT 10

178875

receptor tyrosine kinase - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Aug-2004

C:Accession: I78875

R:Finamore, H.; Kelleher, K.; Morris, G.E.; Bean, K.; Merberg, D.M.; Kriz, R.; Morris, J

Oncogene 8, 2293-2298, 1993

A:Title: Molecular cloning of murine FLT and FLT4.

A:Reference number: I58375; MUID:93330572; PMID:8393164

A:Accession: I78875

A>Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1133 <RES>

A:Cross-references: UNIPROT:P35969; GB:L07297; NID:G293782; PIDN:AAA40078.1; PID:G293783

C:Genetics:

A:Gene: FLT

C:Superfamily: protein kinase homology

C:Keywords: ATP

F:826-1161/Domain: protein kinase homology <KIN>

F:834-842/Region: protein kinase ATP-binding motif

## Query Match

Best Local Similarity 100.0%; Score 52; DB 2; Length 1333;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10

Db 1086 GVLMEIFSL 1095

# RESULT 11

160598

Fit-1 tyrosine kinase receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Aug-2004

C:Accession: I60598

R:Yamane, A.; Seetharam, L.; Yamaguchi, S.; Gotoh, N.; Takahashi, T.; Neufeld, G.; Shibu

Oncogene 9, 2683-2690, 1994

A:Title: A new communication system between hepatocytes and sinusoidal endothelial cells

A:Reference number: I60598; MUID:9433623; PMID:8056332

A:Accession: I60598

A>Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1136 <RES>

A:Cross-references: UNIPROT:P53767; GB:D28498; NID:G511662; PIDN:BAAS0857.1; PID:G600379

C:Superfamily: protein kinase homology

C:Keywords: ATP

F:825-1161/Domain: protein kinase homology <KIN>

F:833-841/Region: protein kinase ATP-binding motif

## Query Match

Best Local Similarity 100.0%; Score 52; DB 2; Length 1336;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10

Db 1086 GVLMEIFSL 1095

# RESULT 12

509982

protein-tyrosine kinase (EC 2.7.1.112) flt1 precursor - human

N:Alternate names: receptor-type tyrosine kinase flt

C:Species: Homo sapiens (man)

C>Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 16-Aug-2004

C:Accession: S09982

R:Shibuya, M.; Yamaguchi, S.; Yamane, A.; Ikeda, T.; Tojo, A.; Matsushima, H.; Sato, M.

Oncogene 5, S19-S24, 1990

A:Title: Nucleotide sequence and expression of a novel human receptor-type tyrosine kina

A:Reference number: S09982; MUID:90221591; PMID:2158038

A:Accession: S09982

A:Molecule type: mRNA

A:Residues: 1-1138 <SHI>

A:Cross-references: UNIPROT:P17948; EMBL:X51602; NID:G31431; PIDN:CAAS5946.1; PID:G31432

C:Genetics:

A:Gene: GDB:FLTI

A:Cross-references: GDB:120616; OMIM:165070

A:Map position: 13q12-13q12

C:Superfamily: protein kinase homology

C:Keywords: ATP; autophosphorylation; glycoprotein; membrane protein; phosphoprotein; ph

F:1-32/Domain: signal sequence #status predicted <SIG>

F:23-1338/Product: protein-tyrosine kinase flt #status predicted <MAT>

F:825-1161/Domain: protein kinase homology <KIN>

F:833-841/Region: protein kinase ATP-binding motif

## Query Match

Best Local Similarity 100.0%; Score 52; DB 2; Length 1338;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10

Db 1086 GVLMEIFSL 1095

# RESULT 13

S51656

vascular endothelial growth factor receptor 1 precursor - Japanese quail

N:Alternate names: quail endothelial kinase 1 (Quel 1); vascular endothelial growth fact

C:Species: Coturnix coturnix japonica (Japanese quail)

C>Date: 07-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 16-Aug-2004

C:Accession: J04953; A56598; I51162; S51656

R:Richmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.M.

Gene 174, 3-8, 1996

A:Title: Molecular cloning of Quel 1 and 2, two quail vascular endothelial growth factor

A:Reference number: J04953; MUID:97017121; PMID:8863722

A:Accession: J04953

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1348 <ETC1>

A:Cross-references: UNIPROT:P52583; EMBL:X63288; NID:G603553; PIDN:CAAS6268.1; PID:G6035

A>Note: submitted to the EMBL Data Library, December 1994

R:Richmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.M.

Mech. Dev. 42, 33-48, 1993

A>Title: Two molecules related to the VEGF receptor are expressed in early endothelial c  
 A:Reference number: A56598; MUID:9337865; PMID:8396413  
 A:Contents: E16 spinal cord  
 A:Accession: A56598  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 910-1348 <E1C2>  
 A:Cross-references: GB:565205; NID:G410680; PIDN:AA828127.1; PID:G410681  
 A>Note: Sequence extracted from NCBI backbone (NCBIT:137162, NCBI:P137163)  
 R:Marcelle, C.; Eichmann, A.  
 Oncogene 7, 2479-2487, 1992  
 A>Title: Molecular cloning of a family of protein kinase genes expressed in the avian em  
 A:Reference number: 150595; MUID:93096482; PMID:1281306  
 A:Accession: 151162  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1023-1079 <MAR>  
 A:Cross-references: EMBL:X69694; NID:G395226; PIDN:CAA49364.1; PID:G938278  
 A>Note: the species is not identified by the authors; the most probable species is shown  
 C:Comment: This protein is an endothelial-specific receptor and binds vascular endothel  
 C:Superfamily: protein kinase homology  
 C:Keywords: ATP; embryo; growth factor receptor; transmembrane protein  
 F.1-19/Domain: signal sequence #status predicted <SIG>  
 F.20-1348/Product: vascular endothelial growth factor receptor 1 #status predicted <MAT>  
 F.1756-777/Domain: transmembrane #status predicted <TM>  
 F.823-1160/Domain: protein kinase homology <KIN>  
 F.831-839/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 1348;  
 Best Local Similarity 100.0%; Pred. No. 0.49; Mismatches 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
 |||||  
 DB 1085 GVLWEIFSL 1094

RESULT 14

protein-tyrosine kinase (EC 2.7.1.112) KDR - human  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Sep-1993 #sequence\_revision 28-Aug-1998 #text\_change 16-Aug-2004  
 C:Accession: J01402; 158357  
 R:Teram, B. I.; Dougher-Vermazen, M.; Carrion, M.E.; Dimitrov, D.; Armellino, D.C.; Gosh  
 Biochem. Biophys. Res. Commun. 187, 1579-1586, 1992  
 A>Title: Identification of the KDR tyrosine kinase as a receptor for vascular endothelia  
 A:Reference number: J01402; MUID:93038635; PMID:11417831  
 A:Accession: J01402  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-48, 'I', 50-616, 'E', 618-700, 'E', 702-748, 'F', 750-1263, 'I', 1265-1290, 'P', 1292-  
 A:Cross-references: UNIPROT:P35968; EMBL:X61656; NID:G31717  
 R:Teram, B. I.; Carrion, M.E.; Kovacs, E.; Rasmussen, B.A.; Eddy, R.L.; Shows, T.B.  
 Oncogene 6, 1677-1683, 1991  
 A>Title: Identification of a new endothelial cell growth factor receptor tyrosine kinase  
 A:Reference number: 158357; MUID:92018935; PMID:1656371  
 A:Accession: 158357  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 3-1356 <TER2>  
 A:Cross-references: GB:104947; NID:G186674; PIDN:AA59459.1; PID:G186675  
 C:Comment: This enzyme acts as a receptor for vascular endothelial cell growth factor.  
 C:Genetics:  
 A:Gene: CDB:KDR  
 A:Cross-references: GDB:127921; OMIM:191306  
 A:Map position: 4q12-4q12  
 C:Superfamily: protein kinase homology  
 C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein ki  
 F.1766-790/Domain: transmembrane #status predicted <TM>  
 F.832-1167/Domain: protein kinase homology <KIN>  
 F.840-848/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 1356;

Best Local Similarity 100.0%; Pred. No. 0.5;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
 |||||  
 DB 1092 GVLWEIFSL 1101

RESULT 15

protein-tyrosine kinase (EC 2.7.1.112) FLT4 - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 16-Aug-2004  
 A:Accession: 158375; B42010  
 R:Finerty, H.; Kelleher, K.; Morris, G.E.; Bean, K.; Merberg, D.M.; Kitz, R.; Morris,  
 Oncogene 8, 2293-2298, 1993  
 A>Title: Molecular cloning of murine FLT and FLT4  
 A:Reference number: 158375; MUID:9330572; PMID:8393164  
 A:Accession: 158375  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1363 <RES>  
 A:Cross-references: UNIPROT:P35917; GB:107296; NID:G293780; PIDN:AA40077.1; PID:G29378  
 R:Galland, F.; Karamysheva, A.; Nattei, M.G.; Rosnet, O.; Marchetto, S.; Birnbaum, D.  
 Genomics 13, 475-478, 1992  
 A>Title: Chromosomal localization of FLT4, a novel receptor-type tyrosine kinase gene.  
 A:Reference number: A42010; MUID:92307693; PMID:1319394  
 A:Accession: B42010  
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
 A:Molecule type: DNA  
 A:Residues: 1033-1072 <GAL>  
 C:Genetics:  
 A:Gene: FLT4  
 C:Superfamily: protein kinase homology  
 C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase  
 F.843-1176/Domain: protein kinase homology <KIN>  
 F.851-859/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 1363;  
 Best Local Similarity 100.0%; Pred. No. 0.5; Mismatches 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
 |||||  
 DB 1101 GVLWEIFSL 1110

Search completed: November 30, 2004, 07:41:23  
 Job time : 21 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 76.4103 Seconds  
(without alignments)  
75.301 Million cell updates/sec

Title: US-10-008-377A-2  
Perfect score: 52  
Sequence: 1 GVLMWELFSL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	52	100.0	09N0K8	09N0K8 sus scrofa
2	52	100.0	028210	028210 bos taurus
3	52	100.0	09X554	09X554 oryctolagus
4	52	100.0	08M123	08M123 ovis aries
5	52	100.0	09N112	09N112 ovis aries
6	52	100.0	09N111	09N111 ovis aries
7	52	100.0	09U8V0	09U8V0 capreolus
8	52	100.0	09U8V8	09U8V8 lampetra re
9	52	100.0	08CD05	08CD05 mus musculu
10	52	100.0	042570	042570 xenopus lae
11	52	100.0	08CUH1	08CUH1 ratcus norv
12	52	100.0	091356	091356 coturnix co
13	52	100.0	09PUL0	09PUL0 brachydantio
14	52	100.0	09CW58	09CW58 mus musculu
15	52	100.0	095Y99	095Y99 halocynthia
16	52	100.0	08TDJ5	08TDJ5 homo sapien
17	52	100.0	08415	08415 homo sapien
18	52	100.0	09GSH3	09GSH3 halocynthia
19	52	100.0	089923	089923 mus musculu
20	52	100.0	08UVR8	08UVR8 fuga rubrip
21	52	100.0	1019	1019
22	52	100.0	1088	1088
23	52	100.0	1161	1161
24	52	100.0	1298	1298
25	52	100.0	1298	1298
26	52	100.0	1327	1327
27	52	100.0	1333	1333
28	52	100.0	1336	1336
29	52	100.0	1338	1338
30	52	100.0	1345	1345
31	52	100.0	1348	1348

32	52	100.0	1356	1	VGR2 HUMAN	P35968 homo sapien
33	52	100.0	1363	2	VGR2 MOUSE	P35917 mus musculu
34	52	100.0	1363	2	086W07	086W07 homo sapien
35	52	100.0	1367	2	0912T1	0912T1 ratcus norv
36	52	100.0	1367	2	VGR2 MOUSE	P35918 mus musculu
37	52	100.0	1379	2	F79701	F79701 coturnix co
38	52	100.0	1620	1	ALK HUMAN	P97701 coturnix co
39	52	100.0	1621	2	ALK MOUSE	P97793 mus musculu
40	51	98.1	143	2	060515	060515 mus musculu
41	51	98.1	207	2	09P8G8	09P8G8 gallus gall
42	51	98.1	345	2	09PVT7	09PVT7 lampetra re
43	51	98.1	395	2	070W10	070W10 clona intes
44	51	98.1	395	2	CAD58833	CAD58833 clona int
45	51	98.1	457	2	Q90269	Q90269 brachydantio

## ALIGNMENTS

RESULT 1  
ID Q9N0K8 PRELIMINARY; PRT; 116 AA.  
AC Q9N0K8;  
DT 01-OCT-2000 (TREMREL.15, Created)  
DT 01-OCT-2000 (TREMREL.15, Last sequence update)  
DT 01-OCT-2003 (TREMREL.25, Last annotation update)  
DE Flt-1 type VEGF receptor (Fragment).  
GN Name=Flt-1;  
OS Sus scrofa (pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sui.  
OX NCBI\_Taxid=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Oviduct;  
RA Kettler A., Mollenhaupt K., Einspanier R.;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ245445; CAB76366.1;  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004713; F:Protein-tyrosine kinase activity; IEA.  
DR GO; GO:0004872; F:Receptor activity; IEA.  
DR GO; GO:0006468; P:Protein amino acid phosphorylation; IEA.  
DR InterPro; IP001109; Kinase like.  
DR InterPro; IP000719; Prot\_Kinase.  
DR InterPro; IP001245; Tyr\_Kinase.  
DR Pfam; PF00069; Pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Prot\_Kinase; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Receptor.  
FT NON\_TER 1  
FT NON\_TER 116  
SQ SEQUENCE 116 AA; 13289 MW; 4663DE769F50FA72 CRC64;  
Query Match 100.0%; Score 52; DB 2; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 1 GVLMWELFSL 10  
DB 9 GVLMWELFSL 18  
RESULT 2  
ID Q28210 PRELIMINARY; PRT; 153 AA.  
AC Q28210;  
DT 01-NOV-1996 (TREMREL.01, Created)  
DT 01-NOV-1996 (TREMREL.01, Last sequence update)  
DT 01-OCT-2003 (TREMREL.25, Last annotation update)  
DE VEGF-Receptor (Flt-Receptor) (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC	Bovinae; Bos.
OX	NCBI_TaxID=9913;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Kidney;
RX	MEDLINE=96212225; PubMed=8626709;
RA	Mandriota S.J., Menoud P., Pepper M.S.;
RT	"Transforming growth factor beta 1 downregulates vascular endothelial
RT	growth factor receptor-2/Flk-1 expression in vascular endothelial
RT	cells."
RL	J. Biol. Chem. 271:11500-11505(1996).
DR	EMBL: X94263; CAA63939.1;
DR	GO: GO:0005524; F:ATP binding; IEA.
DR	GO: GO:0004773; F:Protein-tyrosine kinase activity; IEA.
DR	GO: GO:0004872; F:Receptor activity; IEA.
DR	GO: GO:0006468; P:Protein amino acid phosphorylation; IEA.
DR	InterPro: IPRO1009; Kinase, like.
DR	InterPro: IPRO0719; Prot_kinase.
DR	InterPro: IPRO01245; Tyr_kinase.
DR	Pfam: PF00069; Kinase; 1.
DR	PRINTS: PR00109; TYRKINASE.
DR	ProDom: PD00001; Prot_kinase; 1.
DR	PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KW	Kinase; Receptor; Transferase; Tyrosine-protein kinase.
FT	NON_TER 1
FT	NON_TER 178 178

```

SQ SEQUENCE 178 AA; 20363 MW; B5F1D072A2D0A477 CRC64;
Query Match 100.0%; Score 52; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVLWEIFSL 10
|||||
Db 74 GVLWEIFSL 83

RESULT 4
O8M123
ID O8M123 PRELIMINARY; PRT; 180 AA.
AC O8M123;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Vascular endothelial growth factor receptor-2 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Placental artery endothelium;
RA Chung J.-Y., Teoi S.C.M., Wen Y.-X., Magness R.R., Zheng J.;
RA Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF934634; AN04105.1; -.
DR HSP; P35958; IYR2.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot Kinase.
DR InterPro; IPR01245; Tyr_Kinase.
DR InterPro; IPR008266; Tyr_Kinase_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot Kinase; 1.
DR ProSite; PSS0011; PROTEIN_KINASE_DOM; 1.
DR ProSite; PSS0109; PROTEIN_KINASE_TYR; 1.
KW Kinase; Receptor; Transferase; Tyrosine-protein kinase.
FT NON_TER 1
FT NON_TER 1
FT SEQUENCE 180 AA; 20559 MW; 70E5F444574779A0 CRC64;
SQ SEQUENCE 180 AA; 20559 MW; 70E5F444574779A0 CRC64;
Query Match 100.0%; Score 52; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVLWEIFSL 10
|||||
Db 74 GVLWEIFSL 83

RESULT 5
O9N112
ID O9N112 PRELIMINARY; PRT; 207 AA.
AC O9N112;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE KDR/fik-1 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]

```

SEQUENCE FROM N.A.  
 RA Scholz T.D., Segar J.L.;  
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF233076; AAF60280.1; -  
 DR GO: GO:0005524; P:ATP binding; IEA.  
 DR GO: GO:0004713; P:protein-tyrosine kinase activity; IEA.  
 DR GO: GO:0016740; P:transferase activity; IEA.  
 DR GO: GO:000521; P:vascular endothelial growth factor receptor; IEA.  
 DR GO: GO:000468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase\_like.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase.  
 DR Pfam: PF00069; Kinase; 1.  
 DR PRINTS: PRO0109; TYRKINASE.  
 DR PRODOM: PD000001; Prot\_kinase; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_DOM; 1.  
 DR Kinase; Transferase; Tyrosine-protein kinase.  
 FT NON TER 1  
 FT 207  
 SQ SEQUENCE 207 AA; 23691 MW; 51C3950D82F988FC CRC64;

Query Match 100.0%; Score 52; DB 2; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10  
 Db 81 GVLMWEIFSL 90

|||||  
 81 GVLMWEIFSL 90

## RESULT 6

Q9N11 PRELIMINARY; PRT; 221 AA.

AC Q9N11;  
 DT 01-OCT-2000 (TREMREL. 15, Created)  
 DT 01-OCT-2000 (TREMREL. 15, Last sequence update)  
 DT 01-OCT-2003 (TREMREL. 25, Last annotation update)  
 DE F1t-1 (Fragment).

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Caprinae; Ovis;

OX NCBI\_TaxID=9940;

RN [1]  
 RP SEQUENCE FROM N.A.

RA Scholz T.D., Segar J.L.;  
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF233077; AAF60281.1; -

DR GO: GO:0005524; P:ATP binding; IEA.  
 DR GO: GO:0004713; P:protein-tyrosine kinase activity; IEA.  
 DR GO: GO:0016740; P:transferase activity; IEA.

DR GO: GO:000468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase\_like.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase.

DR Pfam: PF00069; Kinase; 1.  
 DR PRINTS: PRO0109; TYRKINASE.

DR PRODOM: PD000001; Prot\_kinase; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_DOM; 1.  
 DR Kinase; Transferase; Tyrosine-protein kinase.

FT NON TER 1  
 FT 221

SQ SEQUENCE 221 AA; 25356 MW; 5F3059E2C3773B50 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 221;  
 Best Local Similarity 100.0%; Pred. No. 0.33;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10  
 Db 81 GVLMWEIFSL 90

|||||  
 81 GVLMWEIFSL 90

## RESULT 7

Q98V0 PRELIMINARY; PRT; 308 AA.

AC Q98V0;  
 DT 01-MAY-2000 (TREMREL. 13, Created)  
 DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMREL. 25, Last annotation update)  
 DE Kit-like A (Fragment).

OS Epiplatys burgeri (Inshore hapfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Hyperoartii; Myxiniiformes;  
 OC Myxiniidae; Epiplatinae; Epiplatetidae.

OX NCBI\_TaxID=7764;

RN [1]  
 RP SEQUENCE FROM N.A.

RA MEDLINE=20020330; PubMed=10552041;  
 RA Suga H., Hoshiyama D., Kuraku S., Katoh K., Kubokawa K., Miyata T.;

RT "Protein tyrosine kinase CDNA from amphioxus, hapfish, and lamprey;  
 RT isoform duplications around the divergence of cyclostomes and

RT gnathostomes";  
 RT J. Mol. Evol. 49:601-608(1999).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. CSF-

CC 1/PDGF receptor subfamily.

DR EMBL: AB025553; BAA84743.1; -

DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0005524; P:ATP binding; IEA.  
 DR GO: GO:0004714; P:transmembrane receptor protein tyrosine kin.

DR GO: GO:000468; P:protein amino acid phosphorylation; IEA.  
 DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.  
 DR InterPro: IPR011009; Kinase\_like.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR001824; ReceptTykKinaseIII.

DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase.

DR Pfam: PF00069; Kinase; 1.  
 DR PRODOM: PD000001; Prot\_kinase; 1.

DR SMART: SMO0219; TyrcK; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.

DR PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR Immunoglobulin domain; Kinase; Receptor; Transferase; Transmembrane;

KM Tyrosine-protein kinase.  
 FT NON TER 1  
 FT 308

SQ SEQUENCE 308 AA; 34459 MW; FC63206DC6E629B8 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 0.45;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10  
 Db 233 GVLMWEIFSL 242

|||||  
 233 GVLMWEIFSL 242

## RESULT 8

Q9PV08 PRELIMINARY; PRT; 314 AA.

AC Q9PV08;  
 DT 01-MAY-2000 (TREMREL. 13, Created)  
 DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMREL. 25, Last annotation update)  
 DE Kit-like A (Fragment).

OS Lampetra reissneri (Far Eastern brook lamprey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartii;  
 OC Petromyzontiformes; Petromyzontidae; Lethenteron.

OX NCBI\_TaxID=7753;

RN [1]

```

RP      SEQUENCE FROM N.A.
RX      MEDLINE=20020330; PubMed=10552041;
RA      Suga H., Hoshiyama D., Kikaku S., Katch K., Kubokawa K., Miyata T.,
RT      "Protein tyrosine kinase CDNA's from amphioxus, hagfish, and lamprey:
RT      isoform duplications around the divergence of cyclostomes and
RT      teleostomes."
RL      J Mol. Evol. 48:601-608(1999).
DR      EMBL; AB023555; BA84745.1; -.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR      InterPro; IPR011009; Kinase, like.
DR      InterPro; IPR000719; Prot kinase.
DR      InterPro; IPR001245; Tyr kinase.
DR      InterPro; IPR008266; Tyr_kinase_AS.
DR      Pfam; PF00069; Pkinase; Tyr_kinase_AS.
DR      ProDom; PD000001; Prot_kinase; 2.
DR      SMART; SM00219; TYKc; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
RW      Kinase; transferase; Tyrosine-protein kinase.
PT      NON_TER
SQ      SEQUENCE 314 AA; 35254 MW; D375DB5B7D8DDAFC CRC64;

Query March 100.0%; Score 52; DB 2; Length 314;
Best Local Similarity 100.0%; Pred No. 0.46;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps

CY      1 GVLIWTFSL 10
Db      243 GVLIWTFSL 252

RESULT 9
O8CD05 PRELIMINARY; PRT; 341 AA.
ID O8CD05.
AC O8CD05.
DT 01-MAR-2003 (TRMBLpre1. 23; Created)
DT 01-MAR-2003 (TRMBLpre1. 23; Last sequence update)
DE 01-MAR-2004 (TRMBLpre1. 26; Last annotation update)
DE Mus musculus 14 days embryo thymus cDNA, RIKEN full-length enriched
DE library, clone:613041C07 product:kinase insert domain protein
DE receptor, full insert sequence. (Fragment).
DE
DE Name=Rdr;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.,
RT "High-efficiency full-length cDNA cloning.";
RT Mech. Enzymol. 303:19-44(1999).
RN 2
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=1121781;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN 3
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Thymus;
RA The PANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN 4
RP SEQUENCE FROM N.A.

```

RC	STRAIN=C57BL/6J; TISSUE=Thymus;
RX	MEDLINE=204939374; PubMed=11042159;
RA	Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA	Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT	"Normalization and subtraction of cap-trapper-selected cDNAs to
RT	prepare full-length cDNA libraries for rapid discovery of new genes.";
RL	Genome Res. 10:1167-1630(2000).
RN	[5]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Thymus;
RX	MEDLINE=20530913; PubMed=11076861;
RA	Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA	Kono H., Akiyama J., Nishi K., Kitunai T., Tashtiro H., Itoh M.,
RA	Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA	Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasaiwa K.,
RA	Fujiwaka S., Inoue K., Togawa Y., Irawa M., Ohara E., Wachihi M.,
RA	Ozmeda Y., Ichikawa T., Ozawa K., Tanaka T., Matsura S., Kanai C.,
RT	"RIKEN integrated sequence analysis (RISA) system-364-format
RT	sequencing pipeline with 364 multicapillary sequencer.";
RL	Genome Res. 10:1157-1771(2000).
RN	[6]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Thymus;
RA	Adechi U., Aizawa K., Akimaru T., Atakawa T., Bono H., Carninci P.,
RA	Fukuda S., Furuno M., Hanagata T., Hara A., Hashizume W.,
RA	Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA	Hori F., Imetani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA	Kaion H., Kawai T., Kojima Y., Kondo S., Kono H., Kotoda M., Koya S.,
RA	Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA	Nishi K., Nomura K., Numazaki R., Ono M., Osato N., Okazaki Y.,
RA	Saito R., Satoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA	Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA	Tomaru A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA	Taniguchi A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL	Submitted (Jul-2001) to the EMBL/GenBank/DDSI databases.
DR	EMBL: AK031739; BAC27532.1;
DR	HSPG; P35968; 1VR2
DR	MGSD; MG136683; Kd1.
DR	GO: GO:0045165; P:cell fate commitment, IMP.
DR	GO: GO:0045446; P:endothelial cell differentiation, IDA.
DR	Interpro: IPRO11009; Kinase-like.
DR	Interpro: IPRO00719; Prot_Kinase.
DR	Interpro: IPRO1245; Ty_Kinase.
DR	Interpro: IPRO08265; Ty_Kinase_AS.
DR	Interpro: IPRO09136; VEGFR2.
DR	Pfam: PF00069; Pkinase; 1.
DR	PRINTS: PR00109; TYRKINASE.
DR	PRINTS: PRO1834; VEGFRRECEPT2.
DR	ProDom: PD000301; Prot_Kinase; 1.
DR	PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE: PS00109; PROTEIN_KINASE_TYK; 1.
KW	Kinase; Receptor; transferase; tyrosine-protein kinase.
FT	NON TER
Q	SEQUENCE 341 AA; 38302 MW; E2B4DC4BA81195 CRC64;
Query Match	100.0%; Score 52; DB 2; Length 341;
Best Local Similarity	100.0%; Pred. No. 0.5;
Matches 10; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GVLMWFISL 10
D5	
Db	86 GVLMWFISL 95
RESULT 10	
ID	042570
AC	PRELIMINARY;
PT	346 AA.
DT	01-JAN-1998 (TrEMBLrel. 05; Created)
DT	01-OCT-1998 (TrEMBLrel. 05; Last Sequence Update)
DT	01-MAR-2004 (TrEMBLrel. 26; Last Annotation Update)
DB	Receptor tyrosine kinase [Fragment].



```

GN Name=flx-1;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenoidea; Xenopus.
OC NCBI_Taxid=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RA Cleaver O., Tonissen K.F., Saha M.S., Krieg P.A.;
RT "Neovascularization of the Xenopus embryo.";
RL Dev. Dyn. 0:0-0(1997).
DR EMBL; AF007760; F:ATP binding; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004672; F:receptor activity; IEA.
DR GO; GO:0016740; F:transferrase activity; IEA.
DR GO; GO:0005021; F:vascular endothelial growth factor receptor; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR009136; VEGFR2.
DR Pfam; PF00069; Kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRINTS; PR01834; VEGFRECEPTOR2.
DR PRODOM; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
GN Kinase; Receptor; Transferrase; Tyrosine-protein kinase.
FT NON_TER 1
FT NON_TER 346
SQ SEQUENCE 346 AA; 38976 MW; AA35FA8F0C702416 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 346;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10
Db 88 GVLMWEIFSL 97

RESULT 11
Q8CJH1 PRELIMINARY; PRT; 351 AA.
AC Q8CJH1;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Anaplastic lymphoma kinase (Fragment).
GN Name=alk;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RA Asahina K., Sato H., Yamasaki C., Tateno C., Kataoka M., Shiohawa M.,
RA Katayama S., Yoshizato K.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB073169; BAC2163.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferrase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002011; ReceptTyRKinII.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.

```

```

DR Pfam; PF00069; Kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00219; TYRKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; UNKNOWN_1.
DR ATP-binding; Kinase; 1.
FT NON_TER 1
FT NON_TER 351
SQ SEQUENCE 351 AA; 38984 MW; 6ADCEC220C76B125 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10
Db 319 GVLMWEIFSL 328

RESULT 12
Q91356 PRELIMINARY; PRT; 367 AA.
AC Q91356;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Endothelial kinase Quek2 (Fragment).
GN Name=endothelial kinase Quek2;
OS Corumix corumix (Common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Cuculix.
OC NCBI_Taxid=9091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93378866; PubMed=8396413;
RT Eichmann A., Marcelle C., Breat C., le Douarin N.M.;
RT "Two molecules related to the VEGF receptor are expressed in early
RT endothelial cells during avian embryonic development.";
RL Mech. Dev. 42:33-48(1993).
DR EMBL; S65207; AAB28128.1; -.
DR PIR; B56598; B56598.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferrase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00069; Kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
GN Kinase; Transferrase; Tyrosine-protein kinase.
FT NON_TER 367
FT NON_TER 367
SQ SEQUENCE 367 AA; 41792 MW; 04442B35940027A0 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10
Db 101 GVLMWEIFSL 110

RESULT 13
Q9PULO

```

ID Q9PUL0 PRELIMINARY; PRT; 404 AA.  
 AC Q9PUL0  
 DT 01-MAY-2000 (TEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE P14 protein (Fragment).  
 GN Name: P14;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 CC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98294174; PubMed=9630750;  
 RA Thompson M.A., Ransom D.G., Pratt S.J., MacLennan H., Kieran M.W.,  
 RA Derrich H.W., I.L.I., Vail B., Huber T.L., Paw B., Brownlie A.J.,  
 RA Oates A.C., Fritz A., Gates M.A., Amos A., Bahary N., Talbot W.S.,  
 RA Her H., Beler D.R., Postlethwait J.H., Zon L.I.;  
 RT "The cloche and spadetail genes differentially affect hematopoiesis  
 RT and vasculogenesis.";  
 RL Dev. Biol. 197;248-269 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Thompson M.A., Ransom D.G., Pratt S.J., MacLennan H., Kieran M.W.,  
 RA Derrich H.W., I.L.I., Vail B., Huber T.L., Paw B., Brownlie A.J.,  
 RA Oates A.C., Fritz A., Gates M.A., Amos A., Bahary N., Talbot W.S.,  
 RA Her H., Beler D.R., Postlethwait J.H., Zon L.I.;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF178759; AAD56011.2;  
 DR ZFIN; ZDB-GENE-980526-326; flt4.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011009; Kinase\_like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; Kinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD0000001; Prot\_kinase; 1.  
 DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PSS0109; PROTEIN\_KINASE\_TYR; 1.  
 DR KINASE; transferase; Tyrosine-Protein kinase.  
 FT NON\_TER  
 SQ SEQUENCE 404 AA; 45281 MW; A149B836219EE4A9 CRC64;  
 Query March 100.0%; Score 52; DB 2; Length 404;  
 Best Local Similarity 100.0%; Pred. No. 0.59; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GVLLWETPSL 10  
 Db 156 GVLLWETPSL 165  
 RESULT 14  
 Q9CWS8 PRELIMINARY; PRT; 453 AA.  
 AC Q9CWS8  
 DT 01-JUN-2001 (TEMBLrel. 17, Created)  
 DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)  
 DT 01-MAY-2004 (TEMBLrel. 26, Last annotation update)  
 DE Mus musculus adult male lung cDNA, RIKEN full-length enriched library,  
 DE clone:1200012B22 product:FMS-like tyrosine kinase 1, full insert  
 DE sequence. (Fragment).  
 GN Name: Flt1;  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN PANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RA The PANTOM Consortium  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,  
 RA Kono H., Akiyama Y., Nishi K., Kusunagi T., Teshiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
 RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kasaiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawat T.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system: 384-format  
 RT sequencing pipeline with 384 multiplexed sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arikawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,  
 RA Maruyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ono M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK004721; BAB23504.2;  
 DR HSP; P35869.1; ITR2  
 DR MGD; MGI:95858; Flt1.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR001009; Kinase\_like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; Kinase; 2.  
 DR ProDom; PD0000001; Prot\_kinase; 2.  
 DR SMART; SMO0219; TyrKc; 1.

DR PROSITE; PS00109; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 KM Kinase; Transferase; Tyrosine-protein kinase.  
 FT NON TER 1  
 SQ SEQUENCE 453 AA; 51239 MW; 2A48AD0DFC677535 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 453;  
 Best Local Similarity 100.0%; Pred. No. 0.66;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
 |||||  
 Db 206 GVLWEIFSL 215

RESULT 15  
 ID Q95YM9 PRELIMINARY; FRT; 763 AA.  
 AC Q95YM9;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE RGRF.  
 GN Name=Hrfgfr.  
 OS Halocynthia roretzi (Sea squirt).  
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
 OC Stolidobranchia; Pyridae; Halocynthia.  
 OX NCBI\_TaxID=7729;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21417207; PubMed=11526077;  
 RA Shimouchi Y., Murakami S.D., Satoh N.;  
 RT "Fgf signals are involved in the differentiation of notochord cells  
 and mesenchyme cells of the ascidian Halocynthia roretzi.";  
 RL Development 128:2711-2721(2001).  
 DR EMBL; AB046873; BAB59007.1; -.  
 DR HSSP; P11362; 1EVT.  
 DR GO; GO:0005624; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR011009; Kinase\_like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00047; Ig\_2.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR SMART; SM00408; IGC2; 2.  
 DR SMART; SM00219; TYKC; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 2.  
 DR PROSITE; PS00109; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 KM Kinase; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 763 AA; 86741 MW; A34C1871DB796950 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 763;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
 |||||  
 Db 601 GVLWEIFSL 610

Search completed: November 30, 2004, 07:17:27  
 Job time : 77.4103 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12; Search time 77.1795 Seconds  
(Without alignments)  
46,460 Million cell updates/sec

Title: US-10-008-377a-2  
Perfect score: 52  
Sequence: 1 GVLTMEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 2002273 seqs, 358729299 residues  
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: A Genesep 23Sep04:\*  
1: Genesep1980s:\*  
2: Genesep1990s:\*  
3: Genesep2000s:\*  
4: Genesep2001s:\*  
5: Genesep2002s:\*  
6: Genesep2003as:\*  
7: Genesep2003bs:\*  
8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	-6	AAO22984
2	52	100.0	116	4	AAO22984 Human p37
3	52	100.0	116	4	AAO22984 Human ALK
4	52	100.0	116	4	AAO22984 Leukocyte
5	52	100.0	116	6	ABG74477 Tyrosine
6	52	100.0	116	6	ABG74481 Tyrosine
7	52	100.0	127	4	ABBI1041 Human pro
8	52	100.0	316	3	AA87417 Human VEG
9	52	100.0	348	2	AA85938 Human VEG
10	52	100.0	366	3	AA87421 Human VEG
11	52	100.0	525	2	AA87521 Human VEG
12	52	100.0	525	6	ABU04368 Human exp
13	52	100.0	572	2	AA87520 Partial A
14	52	100.0	675	4	ABG20950 Novel hum
15	52	100.0	680	4	AA873858 Human NPM
16	52	100.0	680	5	ABG95067 Human tra
17	52	100.0	680	6	ABU04347 Human exp
18	52	100.0	680	6	ABU04362 Human exp
19	52	100.0	680	6	ABU04358 Human exp
20	52	100.0	713	7	ABG74472 NPM/ALK f
21	52	100.0	713	7	ABG74472 NPM/ALK f
22	52	100.0	771	8	ADOC4223 Human VEG
23	52	100.0	832	8	ADH40881 Human kin
24	52	100.0	1273	3	AA87051 Human kxp
25	52	100.0	1298	2	AA85937 Protein t

26	52	100.0	1298	2	AA85938 FLT4 rece
27	52	100.0	1298	2	ADP90836 Human fms
28	52	100.0	1298	3	AA85935 Human FLT
29	52	100.0	1298	3	AA870746 Human tyr
30	52	100.0	1298	3	AA859542 Mouse act
31	52	100.0	1298	4	AA877855 Human FLT
32	52	100.0	1298	4	AA877855 Human FLT
33	52	100.0	1298	5	ABG32043 Human vas
34	52	100.0	1298	6	ABU07853 Human exp
35	52	100.0	1298	6	ABU04292 Human exp
36	52	100.0	1298	6	ABU04307 Human exp
37	52	100.0	1298	6	ABU04298 Human exp
38	52	100.0	1298	6	ABU04304 Human exp
39	52	100.0	1298	6	ABU04296 Human exp
40	52	100.0	1298	6	ABU04277 Human exp
41	52	100.0	1298	6	ABU04274 Human exp
42	52	100.0	1298	6	ABU04283 Human exp
43	52	100.0	1298	6	ABU04294 Human exp
44	52	100.0	1298	6	ABU04293 Human exp
45	52	100.0	1298	7	ADD08956 Human VEG

## ALIGNMENTS

RESULT 1  
AAO22984 standard; peptide; 10 AA.  
ID AAO22984  
AC AAO22984;  
XX 17-SEP-2003 (first entry)  
DT Human p37-85 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.  
DE HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
XX cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;  
XX oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
KW t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
KW p37-85.  
XX Homo sapiens.  
OS WO2003042243-A2.  
XX 22-MAY-2003.  
XX 14-NOV-2002; 2002NO-EP012764.  
XX 15-NOV-2001; 2001US-0008377.  
XX (NAST-) IST NAZ STUDIO E CURA DEI TUMORI.  
XX (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX Gambacorti-Passerini C, Passoni L;  
XX WPI; 2003-441791/41.  
XX New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
XX useful for preparing a composition for treating ALK-positive lymphoma,  
XX neuroblastoma or ALK-expressing neoplasia.  
PS Claim 1; Page 7; 33pp; English.  
CC The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma  
CC kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a  
CC cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase  
CC which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase  
CC fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell  
CC lines. More than 50% of ALCL cases possess a t(2;5) chromosomal  
CC translocation that leads to the expression of the NPM/ALK fusion protein  
CC which forms a potent oncogene when constitutively activated. Translocated  
CC ALK is a widely expressed tumour-associated antigen characteristic of ALK

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumor cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilized during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p376-85  
 CC ALK-derived H1A-A\*0201 restricted CTL epitope peptide of the invention  
 XX

Sequence 10 AA:

Query Match 100.0%; Score 52; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.017;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
 DB 1 GVLLWEIFSL 10

RESULT 2  
 ID AAB73863 standard; protein; 116 AA.  
 AC AAB73863;  
 XX  
 XX 15-MAY-2001 (first entry)

DE Human ALK catalytic domain sequence #4.

KW Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;  
 KW Chromosomal translocation; cancer; NPM/ALK; catalytic domain.

OS Homo sapiens.

PN US6174674-B1.

PD 16-JAN-2001.

PF 19-JUN-1998; 98US-00100089.

PR 03-DEC-1993; 93US-00160861.

PR 12-OCT-1995; 95US-00542363.

PA (SQUJ-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Morris SW, Look AT;

PI WPI; 2001-243208/25.

PT Detection of chromosomal rearrangement or translocations present in t(2;  
 PT 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 PT techniques.

PS Example 2; Fig 2C; 87pp; English.

CC The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2/5) lymphoma cells. The method is useful for detecting a  
 CC t(2/5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 XX

Sequence 116 AA:

Query Match 100.0%; Score 52; DB 4; Length 116;

Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
 DB 56 GVLLWEIFSL 65

RESULT 3  
 ID AAB73867 standard; protein; 116 AA.  
 AC AAB73867;  
 XX  
 XX 15-MAY-2001 (first entry)

DE Leukocyte tyrosine kinase catalytic domain sequence #4.

KW Leukocyte tyrosine kinase; LTK; protein tyrosine kinase; PTK;  
 KW Insulin receptor subfamily; anaplastic lymphoma kinase; ALK;  
 KW nucleophosmin; NPM; chromosomal translocation; cancer; NPM/ALK;  
 KW catalytic domain.

OS unidentified.

PN US6174674-B1.

PD 16-JAN-2001.

PF 19-JUN-1998; 98US-00100089.

PR 03-DEC-1993; 93US-00160861.

PR 12-OCT-1995; 95US-00542363.

PA (SQUJ-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Morris SW, Look AT;

PI WPI; 2001-243208/25.

PT Detection of chromosomal rearrangement or translocations present in t(2;  
 PT 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 PT techniques.

PS Example 2; Fig 2C; 87pp; English.

CC The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2/5) lymphoma cells. The method is useful for detecting a  
 CC t(2/5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 XX

Sequence 116 AA:

Query Match 100.0%; Score 52; DB 4; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
 DB 56 GVLLWEIFSL 65

RESULT 4  
 ID AAB74477 standard; protein; 116 AA.

```

XX AC ABG74477;
XX XX
XX DT 11-APR-2003 (first entry)
XX XX
XX DE Tyrosine kinase insulin receptor ALK fragment SEQ ID 15.
XX XX
XX KM ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;
XX KM t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;
XX KM nucleolar phosphoprotein; centromeric; telomeric.
XX XX
XX OS Homo sapiens.
XX PN US6451997-B1.
XX PD 17-SEP-2002.
XX PF 28-SEP-2000; 2000US-00670827.
XX PR 03-DEC-1993; 93US-00160861.
XX PR 12-OCT-1995; 95US-00542363.
XX PR 19-JUN-1998; 98US-00100089.
XX PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX PI Morris SW, Look AT;
XX DR WPI; 2003-101739/09.
XX PT Kit for use in method of detecting t(2;5) chromosomal rearrangements or
XX PT rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic
XX PT lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK
XX PT genes.
XX PS Example 1; Fig 2C; 87pp; English.
XX CC This invention describes a novel kit for use in a method of detecting t(2
XX CC ;5) chromosomal rearrangements (CR) or CRs involving nucleolar
XX CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method
XX CC comprises labelled probes that hybridize to a sequence of 2442 or 6226
XX CC nucleotides or their complement, and to region of human chromosome 5/2
XX CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to
XX CC nucleotide). This sequence represents a polypeptide sequence described in
XX CC the disclosure of the invention
XX SQ Sequence 116 AA;
XX
XX Query Match 100.0%; Score 52; DB 6; Length 116;
XX Best Local Similarity 100.0%; Pred. No. 0.22;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GVLWEIFSL 10
XX |||||
XX DB 56 GVLWEIFSL 65
XX
XX RESULT 5
XX ID ABG74481 standard; protein; 116 AA.
XX XX
XX AC ABG74481;
XX XX
XX DT 11-APR-2003 (first entry)
XX XX
XX DE Tyrosine kinase insulin receptor LTK fragment SEQ ID 19.
XX XX
XX KM ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;
XX KM t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;
XX KM nucleolar phosphoprotein; centromeric; telomeric.
XX XX
XX OS Homo sapiens.
XX PN US6451997-B1.

```

```

XX PD 17-SEP-2002.
XX XX
XX PF 28-SEP-2000; 2000US-00670827.
XX XX
XX PR 03-DEC-1993; 93US-00160861.
XX PR 12-OCT-1995; 95US-00542363.
XX PR 19-JUN-1998; 98US-00100089.
XX PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX PI Morris SW, Look AT;
XX DR WPI; 2003-101739/09.
XX PT Kit for use in method of detecting t(2;5) chromosomal rearrangements or
XX PT rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic
XX PT lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK
XX PT genes.
XX PS Example 1; Fig 2C; 87pp; English.
XX CC This invention describes a novel kit for use in a method of detecting t(2
XX CC ;5) chromosomal rearrangements (CR) or CRs involving nucleolar
XX CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method
XX CC comprises labelled probes that hybridize to a sequence of 2442 or 6226
XX CC nucleotides or their complement, and to region of human chromosome 5/2
XX CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to
XX CC nucleotide). This sequence represents a polypeptide sequence described in
XX CC the disclosure of the invention
XX SQ Sequence 116 AA;
XX
XX Query Match 100.0%; Score 52; DB 6; Length 116;
XX Best Local Similarity 100.0%; Pred. No. 0.22;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GVLWEIFSL 10
XX |||||
XX DB 56 GVLWEIFSL 65
XX
XX RESULT 6
XX ID ABB11041 standard; peptide; 127 AA.
XX XX
XX AC ABB11041;
XX XX
XX DT 11-JAN-2002 (first entry)
XX XX
XX DE Human protein-Tyr kinase homologue, SEQ ID NO:1411.
XX XX
XX KM Human; cytokine; cell proliferation; cell differentiation; growth factor;
XX KM haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX KM inhibin; chemokinesis; chemokinesis; thrombolysis; oncogenesis;
XX KM proliferation; metastasis; cancer; tumor; haematopoietic disorder;
XX KM myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX KM chronic inflammatory condition; proliferative retinopathy;
XX KM atherosclerosis; coronary heart disease; arterial ischaemia;
XX KM bone disorder; osteoporosis; vascular growth disorder;
XX KM tissue regeneration; wound healing; infection; immune disorder;
XX KM cell culture; drug screening; gene therapy; antiinflammatory;
XX KM antiasthmatic; antiarthritis; haemostatic; antiatherosclerotic;
XX KM cyostatic; osteopathic; vasotropic; cardiac; virucide; antibacterial;
XX KM antifungal; vulnery; antitumor.
XX XX
XX OS Homo sapiens.
XX PN WO200157188-A2.
XX PD 09-AUG-2001.
XX PF 05-FEB-2001; 2001WO-US003800.

```





Best Local Similarity 100.0%; Pred. No. 0.61;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 237 GVLWEIFSL 246

RESULT 8  
AAR85938  
AAR85938 standard; protein; 348 AA.  
XX AAR85938;  
XX  
XX 14-FEB-1996 (first entry)  
XX Protein tyrosine-kinase SAL-SI.  
XX Protein tyrosine-kinase SAL-SI; cell growth; differentiation.  
XX Homo sapiens.  
XX WO9527061-A1.  
XX 12-OCT-1995.  
XX 04-APR-1995; 95WO-US004228.  
XX 04-APR-1994; 94US-00222616.  
XX (GETH) GENENTECH INC.  
XX Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP, Wood WI;  
XX WPI; 1995-366160/47.  
XX N-PSDB; AAT03101.  
XX  
XX Agonist antibodies which activate specific protein tyrosine kinase(s) -  
XX also activate chimeric proteins of kinase extracellular domain and Ig  
XX constant domain, useful for studying, and therapeutic modulation of, cell  
XX growth and differentiation.  
XX  
XX Disclosure; Page 47-48; 125pp; English.  
XX  
XX DNA probes based on protein tyrosine-kinase (PTK) sequences were used to  
XX screen cDNA libraries to identify novel PTK genes. The SAL-SI gene was  
XX isolated from several megakaryocytic cell line libraries. The gene (see  
XX AAT03101) encodes a protein (AAR85938) showing homology to FLT/FLK family  
XX PTKs. SAL-SI can be used to target proteins in drug design, as an  
XX (ant)agonist of other PTKs, to screen PTK ligands or to modulate  
XX megakaryocyte/platelet adhesion  
XX  
XX Sequence 348 AA;  
SQ

Query Match 100.0%; Score 52; DB 2; Length 348;  
Best Local Similarity 100.0%; Pred. NO. 0.67;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 151 GVLWEIFSL 160

RESULT 9  
AAY87421  
AAY87421 standard; protein; 366 AA.  
XX AAY87421;  
XX  
XX 03-JUL-2000 (first entry)  
XX  
XX Human VEGFR-2 catalytic kinase domain.  
DE  
XX

KM VEGFR-2; vascular endothelial growth factor receptor-2; angiogenesis;  
KM cytoplasmic domain; kinase insert domain; KID; catalytic region;  
KM receptor tyrosine kinase; X-ray crystallography; drug design.

OS Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
XX Domain 128..195  
XX /note="Kinase insert domain (KID)"  
XX  
XX WO200014105-A1.  
XX  
XX 16-MAR-2000.  
XX  
XX 07-SEP-1999; 99WC-US020319.  
XX  
XX 08-SEP-1998; 98US-0099503P.  
XX  
XX (AGOU-) AGOURN PHARM INC.  
XX  
XX McFigue MA, Panko C, Parast CV, Gehring MR, Kan C, Appel K;  
XX Wickersham JA, Showalter RE, Tempczyk-Russell A, Mroczkowski B;  
XX Villalanza JE;  
XX WPI; 2000-256936/22.  
XX  
XX Assaying a candidate compound for its ability to interact with a modified  
XX receptor tyrosine kinase polypeptide, especially vascular endothelial  
XX growth factor-2, by using molecular modelling techniques such as x-ray  
XX crystallography.  
XX  
XX Example 1; Fig 1; 47pp; English.

The invention relates to methods of assaying candidate drugs for their  
ability to interact with modified receptor tyrosine kinases (RTKs) via  
cloning, sequencing and X-ray crystallography. In particular, it relates  
to the identification of vascular endothelial growth factor receptor-2  
(VEGFR-2) agonists and antagonists via the use of X-ray crystallographic  
data obtained from a mutant human VEGFR-2 catalytic cytoplasmic domain.  
VEGFR-2 50P (AAY87421). The cytoplasmic domain of VEGFR-2, a key enzyme  
in angiogenesis, has receptor tyrosine kinase (RTK) activity and mediates  
signal transduction. VEGFR-2 50P comprises residues 806-939 and 990-1171  
of wild-type human VEGFR-2. Relative to the wild-type VEGFR-2, VEGFR-2  
50P sequence has a deletion of 50 amino acids in the kinase insert domain  
(KID), retaining only 18 amino acids of the KID. The KID is  
characteristic of the platelet-derived growth factor receptor (PDGFR)  
family, to which VEGFR-2 belongs. It is not thought to be necessary for  
intrinsic tyrosine kinase activity, but is important for the binding of  
other proteins involved in signal transduction via autophosphorylation of  
KID tyrosine residues. VEGFR-2 50P retains comparable in vitro RTK  
activity to that of wild-type VEGFR-2 kinase domain (AAY87421), but is  
crystallizable, allowing the structure of the phosphorylated unliganded  
VEGFR-2 50P to be determined via X-ray crystallography to a resolution of  
2.4 Angstroms. The structure of the VEGFR-2 50P can provide insights into  
the orientation of the KID of VEGFR-2, which may be relevant to other  
PDGFR family members. The structural information can be used in molecular  
modelling to model the interaction of the VEGFR-2 catalytic domain with  
candidate agonists and antagonists, and can be used in de novo drug  
design. Compounds identified via the methods of the invention can be used  
for treating or preventing diabetic retinopathy or Kaposi's sarcoma,  
suppressing the development of kidney diseases, preventing the joint  
destruction that accompanies rheumatoid arthritis, suppressing  
manifestations of psoriasis, suppressing accelerated atherosclerosis,  
suppressing coronary artery disease and other vasculopathies due to  
atherogenesis, and suppressing tumour growth via paracrine or autocrine  
mediated responses to other cytokines. The present sequence represents  
the wild-type human VEGFR-2 catalytic domain

SQ Sequence 366 AA;  
XX

Query Match 100.0%; Score 52; DB 3; Length 366;  
Best Local Similarity 100.0%; Pred. NO. 0.71;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10  
 Db 287 GVLMWEIFSL 296

## RESULT 10

AA78521  
 ID AAR78521 standard; protein; 525 AA.

XX AAR78521;

XX 25-MAR-2003 (revised)

DT 15-NOV-1995 (first entry)

XX NPM/ALK fusion protein.

XX ALK protein, protein-tyrosine-kinase; fusion protein; NPM protein;  
 KW nucleolar phosphoprotein; anaplastic large cell lymphoma; t(2;  
 KW 5) translocation; diagnosis; gene therapy.

XX Homo sapiens.

OS WO9515331-A1.

XX 08-JUN-1995.

XX 05-DEC-1994; 94WO-US013947.

XX 03-DEC-1993; 93US-00160861.

XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Morris SW, Look AT;

XX WPI; 1995-215226/28.

XX N-PSDB; AAQ95542.

XX Methods for detecting human t(2;5) lymphoma - for detection and diagnosis

XX of anaplastic large cell lymphoma(s).

XX Claim 16; Page 42-44; 70pp; English.

XX The translocation event that occurs in human t(2;5) lymphoma brings  
 CC sequences from the nucleolar phosphoprotein (NPM) gene on chromosome 5q35  
 CC to those from a protein-tyrosine-kinase (ALK) gene on chromosome-2q23.  
 CC Nucleic acids encoding a partial ALK protein (AAR78520) and the ALK/NPM  
 CC fusion protein (AAR78521) were isolated (AAQ95541, AAQ95542).  
 CC Identification of the NPM/ALK fusion allows lymphoma diagnosis, or  
 CC therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR  
 CC 2003 to correct PN field.)

CC Sequence 525 AA;

Qy Query Match 100.0%; Score 52; DB 2; Length 525;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10  
 Db 376 GVLMWEIFSL 385

## RESULT 11

ABU04368  
 ID ABU04368 standard; protein; 525 AA.

XX ABU04368;

XX 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1034.

KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

OS Homo sapiens.

PN WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0356780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOG INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.

XX Example 2; SEQ ID NO 1034; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 525 AA;

Qy Query Match 100.0%; Score 52; DB 6; Length 525;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10  
 Db 376 GVLMWEIFSL 385

## RESULT 12

AA78520  
 ID AAR78520 standard; protein; 572 AA.

XX AAR78520;

XX 25-MAR-2003 (revised)

XX 15-NOV-1995 (first entry)

XX Partial ALK protein.

```

XX ALK protein; protein-tyrosine-kinase; fusion protein; NPM protein;
KM nucleolar phosphoprotein; anaplastic large cell lymphoma; t(2;
KW 5) translocation; diagnosis; gene therapy.
XX Homo sapiens.
XX MO9515331-A1.
XX PD 08-JUN-1995.
XX PF 05-DEC-1994; 94WO-US013947.
XX PR 03-DEC-1993; 93US-00160861.
XX PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX PI Morris SW, Look AT;
XX WPI; 1995-215226/28.
XX DR N-PSDB; AAQ95541.
XX PT Methods for detecting human t(2;5) lymphoma - for detection and diagnosis
XX of anaplastic large cell lymphoma(s).
XX PS Claim 21; Page 40-42; 70pp; English.
XX The translocation event that occurs in human t(2;5) lymphoma brings
CC sequences from the nucleolar phosphoprotein (NPM) gene on chromosome 5q35
CC to those from a protein-tyrosine-kinase (ALK) gene on chromosome 2q32.
CC Nucleic acids encoding a partial ALK protein (AAK8520) and the ALK/NPM
CC fusion protein (AAK78521) were isolated (AAQ95541, AAQ95542).
CC Identification of the NPM/ALK fusion allows lymphoma diagnosis, or
CC therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR
CC -2003 to correct PN field.)
XX SQ Sequence 572 AA;
XX
XX Query Match 100.0%; Score 52; DB 2; Length 572;
XX Best Local Similarity 100.0%; Pred. No. 1.1;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GVILMEIFSL 10
XX 423 GVILMEIFSL 432
XX
XX RESULT 13
XX AAG20950
XX ID AAG20950 standard; protein; 675 AA.
XX AC AAG20950;
XX DT 13-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #20941.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN MO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX

```

```

PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-613962/73.
XX DR N-PSDB; AAB85137.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 51309; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptides and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABB00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pat_sequences
XX SQ Sequence 675 AA;
XX
XX Query Match 100.0%; Score 52; DB 4; Length 675;
XX Best Local Similarity 100.0%; Pred. No. 1.3;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GVILMEIFSL 10
XX 371 GVILMEIFSL 380
XX
XX RESULT 14
XX AAB73858
XX ID AAB73858 standard; protein; 680 AA.
XX AC AAB73858;
XX DT 15-MAY-2001 (first entry)
XX DE Human NPM/ALK fusion protein.
XX KW Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;
XX chromosomal translocation; cancer; NPM/ALK; fusion.
XX OS Homo sapiens.
XX PN US6174674-B1.
XX PD 16-JAN-2001.
XX PF 19-JUN-1998; 98US-00100089.
XX PR 03-DEC-1993; 93US-00160861.
XX PR 12-OCT-1995; 95US-00542363.
XX PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX PI Morris SW, Look AT;
XX

```

DR WPI, 2001-243208/25.  
 DR N-PSDB; AAF76868.  
 XX  
 PT Detection of chromosomal rearrangement or translocations present in t(2;  
 PT 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 PT techniques.  
 XX  
 PS Claim 1, Fig 2a, 87pp; English.  
 CC  
 CC The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 CC  
 SQ Sequence 680 AA;  
 Query Match 100.0%; Score 52; DB 4; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GVLLWEIFSL 10  
 |||||  
 DB 376 GVLLWEIFSL 385  
 RESULT 15  
 ABG95067  
 ID ABG95067 standard; protein; 680 AA.  
 AC  
 AC ABG95067;  
 DT 04-DEC-2002 (first entry)  
 XX  
 DE Human translocation (2; 5) (p23; q35) protein.  
 XX  
 KW Chromosome aberration; oncogenic fusion protein; cancer;  
 KW proliferative disease; cellular protein isoform; heat shock protein 90;  
 KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;  
 KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;  
 KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;  
 KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;  
 KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;  
 KW rhabdomyosarcoma; synovial sarcoma; viral infection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200269900-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 01-MAR-2002; 2002W0-US006518.  
 XX  
 PR 01-MAR-2001; 2001US-0272751P.  
 XX  
 PA (CONF-) CONFORMA THERAPEUTICS CORP.  
 XX  
 PI Fritz LC, Burrows FJ;  
 XX  
 DR WPI; 2002-698710/75.  
 DR N-PSDB; ABB73246.  
 XX  
 PT Treating genetically-defined disease associated with chromosomal  
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative  
 PT diseases, involves administering an inhibitor of heat shock protein 90.

PS Disclosure; Page 200-202; 389pp; English.  
 XX  
 CC The invention describes a method of treating genetically-defined disease  
 CC associated with chromosomal aberrations yielding oncogenic fusion  
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous  
 CC cell population, treating proliferative diseases associated with mutant  
 CC protein or cellular protein isoforms (II) dependent on heat shock protein  
 CC (HSP)-90, or selectively treating cells expressing (II) involving  
 CC administering HSP90-inhibitor. The method is useful for treating  
 CC genetically-defined disease with chromosomal aberration yielding  
 CC oncogenic fusion protein, treating cancerous cells containing fusion  
 CC protein in heterogeneous cell population, treating proliferative disease  
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or  
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.  
 CC p53), or selectively treating cells expressing mutant protein or cellular  
 CC protein isoform in a patient heterozygous for (II). The method is useful  
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell  
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CMML,  
 CC or a disease characterised by a solid tumour such as papillary thyroid  
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and  
 CC synovial sarcoma. The method is also useful for treating viral  
 CC infections. This represents a protein encoded by the DNA sequence of a  
 CC chromosome aberration  
 CC  
 SQ Sequence 680 AA;  
 Query Match 100.0%; Score 52; DB 5; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GVLLWEIFSL 10  
 |||||  
 DB 376 GVLLWEIFSL 385

Search completed: November 30, 2004, 07:39:55  
 Job time : 79.4295 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: November 30, 2004, 07:19:22 : Search time 376.923 Seconds  
(without alignments)  
9.410 Million cell updates/sec

Title: US-10-008-377A-2  
Perfect score: 52  
Sequence: 1 GVLMWEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	10	US-10-008-377-2	Sequence 2, Appl1
2	52	100.0	116	US-09-827-949-15	Sequence 15, Appl1
3	52	100.0	116	US-09-827-949-15	Sequence 19, Appl1
4	52	100.0	127	US-10-276-774-1411	Sequence 1411, Appl1
5	52	100.0	290	US-10-763-418-28	Sequence 28, Appl1
6	52	100.0	291	US-10-763-418-28	Sequence 28, Appl1
7	52	100.0	294	US-10-763-418-17	Sequence 17, Appl1
8	52	100.0	299	US-10-334-143-002	Sequence 202, Appl1
9	52	100.0	317	US-09-939-754-5	Sequence 5, Appl1
10	52	100.0	317	US-09-939-833-5	Sequence 5, Appl1
11	52	100.0	317	US-09-939-833-5	Sequence 5, Appl1
12	52	100.0	348	US-09-982-610-18	Sequence 18, Appl1
13	52	100.0	366	US-10-664-421-26	Sequence 26, Appl1

14	52	100.0	367	US-09-939-754-9	Sequence 9, Appl1
15	52	100.0	367	US-09-939-754-12	Sequence 12, Appl1
16	52	100.0	367	US-09-939-833-9	Sequence 9, Appl1
17	52	100.0	367	US-09-939-833-12	Sequence 12, Appl1
18	52	100.0	367	US-09-939-833-9	Sequence 9, Appl1
19	52	100.0	367	US-09-939-833-12	Sequence 12, Appl1
20	52	100.0	370	US-10-664-421-70	Sequence 70, Appl1
21	52	100.0	429	US-10-664-421-44	Sequence 44, Appl1
22	52	100.0	567	US-10-327-414-8	Sequence 8, Appl1
23	52	100.0	680	US-09-827-949-4	Sequence 4, Appl1
24	52	100.0	1298	US-09-982-610-33	Sequence 33, Appl1
25	52	100.0	1298	US-09-765-534B-2	Sequence 2, Appl1
26	52	100.0	1298	US-10-262-538-12	Sequence 12, Appl1
27	52	100.0	1298	US-10-165-133A-12	Sequence 12, Appl1
28	52	100.0	1298	US-10-394-332A-67	Sequence 67, Appl1
29	52	100.0	1298	US-10-325-048-4	Sequence 4, Appl1
30	52	100.0	1298	US-10-774-802-2	Sequence 2, Appl1
31	52	100.0	1298	US-10-669-176-32	Sequence 32, Appl1
32	52	100.0	1311	US-09-765-534B-5	Sequence 5, Appl1
33	52	100.0	1311	US-10-774-802-5	Sequence 5, Appl1
34	52	100.0	1333	US-10-394-332A-65	Sequence 65, Appl1
35	52	100.0	1338	US-10-060-065-23	Sequence 23, Appl1
36	52	100.0	1338	US-10-059-585-44	Sequence 44, Appl1
37	52	100.0	1338	US-10-021-660-64	Sequence 64, Appl1
38	52	100.0	1338	US-10-090-183-4	Sequence 4, Appl1
39	52	100.0	1338	US-10-116-275-129	Sequence 129, Appl1
40	52	100.0	1338	US-10-211-462-215	Sequence 215, Appl1
41	52	100.0	1338	US-10-621-116-26	Sequence 26, Appl1
42	52	100.0	1339	US-10-165-133A-9	Sequence 9, Appl1
43	52	100.0	1345	US-10-090-183-6	Sequence 6, Appl1
44	52	100.0	1354	US-10-262-538-30	Sequence 30, Appl1
45	52	100.0	1354	US-10-669-176-30	Sequence 30, Appl1

## ALIGNMENTS

RESULT 1  
US-10-008-377-2  
Sequence 2, Application US/10008377  
Publication No. US20030157101A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
APPLICANT: Passoni, Lorena  
TITLE OF INVENTION: Immunogenic ALK Peptides  
FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008,377  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-2

Query Match 100.0%; Score 52; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.031; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0

QY 1 GVLMWEIFSL 10  
DB 1 GVLMWEIFSL 10

RESULT 2  
US-09-827-949-15  
Sequence 15, Application US/09827949  
Patent No. US20010021505A1  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof

FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/670,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 15  
LENGTH: 116  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-15

Query Match 100.0%; Score 52; DB 9; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
|||||  
DB 56 GVLMWEIFSL 65

RESULT 3  
US-09-827-949-19  
Sequence 19, Application US/09827949  
Patent No. US20010021505A1  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/670,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 19  
LENGTH: 116  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-19

Query Match 100.0%; Score 52; DB 9; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
|||||  
DB 56 GVLMWEIFSL 65

RESULT 4  
US-10-276-774-1411  
Sequence 1411, Application US/10276774  
Publication No. US20040053245A1  
GENERAL INFORMATION:  
APPLICANT: Hysaeq, Inc.  
APPLICANT: Tang, Y, Tom et al  
TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides  
FILE REFERENCE: 21272-030

CURRENT APPLICATION NUMBER: US/10/276,774  
CURRENT FILING DATE: 2002-11-18  
PRIOR APPLICATION NUMBER: 09/560,875  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 08/496,914  
PRIOR FILING DATE: 2000-02-03  
NUMBER OF SEQ ID NOS: 2700  
SOFTWARE: Custom  
SEQ ID NO 1411  
LENGTH: 127  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-276-774-1411

Query Match 100.0%; Score 52; DB 15; Length 127;  
Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
|||||  
DB 108 GVLMWEIFSL 117

RESULT 5  
US-10-763-418-28  
Sequence 28, Application US/10763418  
Publication No. US20040185547A1  
GENERAL INFORMATION:  
APPLICANT: MOHAMMADI, MOOSA  
APPLICANT: SCHLESINGER, JOSEPH  
APPLICANT: HUBBARD, STEVAN R.  
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN  
FILE REFERENCE: 038602/0847  
CURRENT APPLICATION NUMBER: US/10/763,418  
CURRENT FILING DATE: 2004-01-26  
PRIOR APPLICATION NUMBER: US/09/664,526  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: 09/188,809  
PRIOR FILING DATE: 1998-11-09  
PRIOR APPLICATION NUMBER: 08/701,191  
PRIOR FILING DATE: 1996-08-21  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 28  
LENGTH: 290  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-763-418-28

Query Match 100.0%; Score 52; DB 16; Length 290;  
Best Local Similarity 100.0%; Pred. No. 0.77;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
|||||  
DB 214 GVLMWEIFSL 223

RESULT 6  
US-10-377-268-28  
Sequence 28, Application US/10377268  
Publication No. US20040171062A1  
GENERAL INFORMATION:  
APPLICANT: HIRTH, KLAUS-PETER  
APPLICANT: MILBURN, MICHAEL VANCE  
TITLE OF INVENTION: METHOD FOR THE DESIGN OF MOLECULAR SCAFFOLDS AND LIGANDS  
FILE REFERENCE: 03963/0303  
CURRENT APPLICATION NUMBER: US/10/377,268  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: 60/437,929  
PRIOR FILING DATE: 2003-01-02  
PRIOR APPLICATION NUMBER: 60/360,651

PRIOR FILING DATE: 2002-02-28  
PRIOR APPLICATION NUMBER: 60/411,398  
PRIOR FILING DATE: 2002-09-16  
PRIOR APPLICATION NUMBER: 60/412,341  
PRIOR FILING DATE: 2002-09-20  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 28  
LENGTH: 291  
TYPE: PRT  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: KDR tyrosine kinase  
US-10-377-268-28

Query Match  
Best Local Similarity 100.0%; Score 52; DB 16; Length 291;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
Db 215 GVLWEIFSL 224

RESULT 7  
US-10-763-418-17  
Sequence 17, Application US/10763418  
Publication No. US20040185547A1  
GENERAL INFORMATION:  
APPLICANT: MOHAMMADI, MOOSA  
APPLICANT: SCHLESSINGER, JOSEPH  
APPLICANT: HUBBARD, STEVAN R.  
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN  
TITLE OF INVENTION: RECEPTOR TYROSINE KINASES  
FILE REFERENCE: 038602/0847  
CURRENT FILING DATE: US/10/763,418  
PRIOR APPLICATION NUMBER: 2004-01-26  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: 09/186,809  
PRIOR FILING DATE: 1998-11-09  
PRIOR APPLICATION NUMBER: 08/701,191  
PRIOR FILING DATE: 1996-08-21  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 17  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-763-418-17

Query Match  
Best Local Similarity 100.0%; Score 52; DB 16; Length 294;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
Db 217 GVLWEIFSL 226

RESULT 8  
US-10-334-143-202  
Sequence 202, Application US/10334143  
Publication No. US20040009549A1  
GENERAL INFORMATION:  
APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
APPLICANT: SUDASANAM, SUCHA  
TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
FILE REFERENCE: 038602/1543  
CURRENT APPLICATION NUMBER: US/10/334,143  
CURRENT FILING DATE: 2002-12-31

PRIOR APPLICATION NUMBER: 60/343,169  
PRIOR FILING DATE: 2001-12-31  
NUMBER OF SEQ ID NOS: 207  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 202  
LENGTH: 299  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (190)  
OTHER INFORMATION: Variable amino acid  
US-10-334-143-202

Query Match  
Best Local Similarity 100.0%; Score 52; DB 15; Length 299;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
Db 223 GVLWEIFSL 232

RESULT 9  
US-09-939-754-5  
Sequence 5, Application US/09939754  
Patent No. US20020051965A1  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPICZYK-RUSSEL, ANNA  
APPLICANT: GERING, MICHAEL R.  
APPLICANT: WROCKZKOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
TITLE OF INVENTION: METHODS OF USE  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/939,754  
CURRENT FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: 09/390,326  
PRIOR FILING DATE: 1993-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-939-754-5

Query Match  
Best Local Similarity 100.0%; Score 52; DB 9; Length 317;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
Db 238 GVLWEIFSL 247

RESULT 10  
US-09-939-832-5  
Sequence 5, Application US/09939832  
Patent No. US20020127538A1  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.

APPLICANT: TEMPCZYK-RUSSEL, ANNA  
APPLICANT: GEHRING, MICHAEL R.  
APPLICANT: MROCKOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/939,832  
CURRENT FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: 09/390,326  
PRIOR FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-939-832-5

Query Match 100.0%; Score 52; DB 9; Length 317;  
Best Local Similarity 100.0%; Pred. No. 0.84;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 238 GVLWEIFSL 247

RESULT 11  
US-09-939-833-5  
Sequence 5, Application US/0939833  
Patent No. US2002016441A1  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPCZYK-RUSSEL, ANNA  
APPLICANT: GEHRING, MICHAEL R.  
APPLICANT: MROCKOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/939,833  
CURRENT FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: US/09/390,326  
PRIOR FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-939-833-5

Query Match 100.0%; Score 52; DB 9; Length 317;  
Best Local Similarity 100.0%; Pred. No. 0.84;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 238 GVLWEIFSL 247

RESULT 12  
US-09-982-610-18  
Sequence 18, Application US/09982610

Patent No. US2002016420A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

Bennett, Brian D.

Goedel, David

Lee, James M.

Matthews, William

Tsai, Siao Ping

Wood, William I.

TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/982,610

FILING DATE: 17-Oct-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/446,648

FILING DATE: 1996-MAY-23

APPLICATION NUMBER: 08/222,616

FILING DATE: 04-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER: 40,378

REFERENCE/DOCKET NUMBER: P0821P3PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994

TELEFAX: 415/852-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Query Match 100.0%; Score 52; DB 9; Length 348;  
Best Local Similarity 100.0%; Pred. No. 0.91;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 151 GVLWEIFSL 160

RESULT 13

US-10-664-421-26

Sequence 26, Application US/10664421

Publication No. US20040142864A1

GENERAL INFORMATION:

APPLICANT: BREWER, RYAN

APPLICANT: IBRAHIM, PRABHA

APPLICANT: KUMAR, ABHINAV

APPLICANT: MANDIVAN, VALSAN V.

APPLICANT: MILBURN, MICHAEL V.

TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE

FILE REFERENCE: 039363/0703

CURRENT APPLICATION NUMBER: US/10/664,421

PRIOR FILING DATE: 2003-09-16

PRIOR APPLICATION NUMBER: 60/412,341



PRIOR APPLICATION NUMBER: 60/411,398  
 PRIOR FILING DATE: 2002-09-16  
 NUMBER OF SEQ ID NOS: 169  
 SOFTWARE: PatentIn Ver. 3.2  
 SEQ ID NO 26  
 LENGTH: 366  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-664-421-26

Query Match 100.0%; Score 52; DB 16; Length 366;  
 Best Local Similarity 100.0%; Pred. No. 0.96;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
 |||||  
 DB 287 GVLWEIFSL 297

RESULT 14  
 US-09-939-754-9  
 Sequence 9, Application US/09939754  
 Patent No. US20020051965A1  
 GENERAL INFORMATION:  
 APPLICANT: MCTIGUE, MICHELE A.  
 APPLICANT: WICKERSHAM, JOHN A.  
 APPLICANT: PINKO, CHRIS  
 APPLICANT: SHOWALTER, RICHARD  
 APPLICANT: PARAST, CAMRAN V.  
 APPLICANT: TEMPCZYK-RUSSEL, ANNA  
 APPLICANT: GEHRING, MICHAEL R.  
 APPLICANT: MROCKZOWSKI, BARBARA  
 APPLICANT: KAN, CHEN-CHEN  
 APPLICANT: VILAFRANCA, J. ERNEST  
 APPLICANT: APPELT, KRZYSZTOF  
 TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
 FILE REFERENCE: 0125-0016US  
 CURRENT APPLICATION NUMBER: US/09/939,754  
 CURRENT FILING DATE: 2001-08-28  
 PRIOR APPLICATION NUMBER: 09/390,326  
 PRIOR FILING DATE: 1999-09-07  
 NUMBER OF SEQ ID NOS: 12  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 9  
 LENGTH: 367  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-939-754-9

Query Match 100.0%; Score 52; DB 9; Length 367;  
 Best Local Similarity 100.0%; Pred. No. 0.96;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
 |||||  
 DB 288 GVLWEIFSL 297

RESULT 15  
 US-09-939-754-12  
 Sequence 12, Application US/09939754  
 Patent No. US20020051965A1  
 GENERAL INFORMATION:  
 APPLICANT: MCTIGUE, MICHELE A.  
 APPLICANT: WICKERSHAM, JOHN A.  
 APPLICANT: PINKO, CHRIS  
 APPLICANT: SHOWALTER, RICHARD  
 APPLICANT: PARAST, CAMRAN V.  
 APPLICANT: TEMPCZYK-RUSSEL, ANNA  
 APPLICANT: GEHRING, MICHAEL R.  
 APPLICANT: MROCKZOWSKI, BARBARA  
 APPLICANT: KAN, CHEN-CHEN

APPLICANT: VILAFRANCA, J. ERNEST  
 APPLICANT: APPELT, KRZYSZTOF  
 TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
 FILE REFERENCE: 0125-0016US  
 CURRENT APPLICATION NUMBER: US/09/939,754  
 CURRENT FILING DATE: 2001-08-28  
 PRIOR APPLICATION NUMBER: 09/390,326  
 PRIOR FILING DATE: 1999-09-07  
 NUMBER OF SEQ ID NOS: 12  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 12  
 LENGTH: 367  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-939-754-12

Query Match 100.0%; Score 52; DB 9; Length 367;  
 Best Local Similarity 100.0%; Pred. No. 0.96;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
 |||||  
 DB 288 GVLWEIFSL 297

Search completed: November 30, 2004, 08:06:35  
 Job time: 377.923 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 25.1282 Seconds  
(without alignments)  
26.392 Million cell updates/sec

Title: US-10-008-377A-2  
Perfect score: 52  
Sequence: 1 GVLMWEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	116	1 US-08-542-363-15	Sequence 15, Appl
2	52	100.0	116	1 US-08-542-363-19	Sequence 19, Appl
3	52	100.0	116	3 US-09-100-089-15	Sequence 15, Appl
4	52	100.0	116	3 US-09-100-089-19	Sequence 19, Appl
5	52	100.0	116	4 US-09-670-827-15	Sequence 15, Appl
6	52	100.0	116	4 US-09-670-827-19	Sequence 19, Appl
7	52	100.0	116	4 US-08-827-949-15	Sequence 15, Appl
8	52	100.0	116	4 US-08-827-949-19	Sequence 19, Appl
9	52	100.0	290	2 US-08-701-91A-28	Sequence 28, Appl
10	52	100.0	290	4 US-09-664-526-28	Sequence 28, Appl
11	52	100.0	294	2 US-08-701-91A-17	Sequence 17, Appl
12	52	100.0	294	4 US-09-664-526-17	Sequence 17, Appl
13	52	100.0	317	3 US-09-390-326-5	Sequence 5, Appl
14	52	100.0	317	4 US-09-390-326-5	Sequence 5, Appl
15	52	100.0	317	4 US-09-939-833-5	Sequence 5, Appl
16	52	100.0	317	4 US-09-939-833-5	Sequence 5, Appl
17	52	100.0	317	4 US-09-939-832-5	Sequence 5, Appl
18	52	100.0	348	1 US-08-222-616-18	Sequence 18, Appl
19	52	100.0	348	4 US-08-446-648-18	Sequence 18, Appl
20	52	100.0	348	5 US-09-982-610-18	Sequence 18, Appl
21	52	100.0	367	3 US-09-390-326-9	Sequence 9, Appl
22	52	100.0	367	3 US-09-390-326-12	Sequence 12, Appl
23	52	100.0	367	4 US-09-939-833-9	Sequence 9, Appl
24	52	100.0	367	4 US-09-939-833-12	Sequence 12, Appl
25	52	100.0	367	4 US-09-506-506-9	Sequence 9, Appl
26	52	100.0	367	4 US-09-506-506-12	Sequence 12, Appl
27	52	100.0	367	4 US-09-939-832-9	Sequence 9, Appl

28	52	100.0	367	4 US-09-939-832-12	Sequence 12, Appl
29	52	100.0	525	1 US-08-160-861-4	Sequence 4, Appl
30	52	100.0	572	1 US-08-160-861-3	Sequence 3, Appl
31	52	100.0	680	1 US-08-542-363-4	Sequence 4, Appl
32	52	100.0	680	4 US-09-100-089-4	Sequence 4, Appl
33	52	100.0	680	4 US-09-670-827-4	Sequence 4, Appl
34	52	100.0	680	4 US-09-827-949-4	Sequence 4, Appl
35	52	100.0	1298	1 US-08-222-616-33	Sequence 33, Appl
36	52	100.0	1298	1 US-08-340-011-2	Sequence 2, Appl
37	52	100.0	1298	3 US-08-901-710-2	Sequence 2, Appl
38	52	100.0	1298	3 US-08-446-648-33	Sequence 33, Appl
39	52	100.0	1298	4 US-09-982-610-33	Sequence 33, Appl
40	52	100.0	1298	5 PCT-US95-04228-33	Sequence 33, Appl
41	52	100.0	1311	1 US-08-340-011-5	Sequence 5, Appl
42	52	100.0	1311	3 US-08-901-710-5	Sequence 5, Appl
43	52	100.0	1338	3 US-08-750-141A-3	Sequence 3, Appl
44	52	100.0	1338	4 US-09-119-014D-6	Sequence 6, Appl
45	52	100.0	1356	1 US-08-810-116-8	Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-08-542-363-15  
Sequence 15, Application US/08542363  
Parent No. 5770421  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: Alk Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,363  
FILING DATE: 12-OCT-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400001/SLF/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-542-363-15

Query Match 100.0%; Score 52; DB 1; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052; Mismatches 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 1 GVLMWEIFSL 10  
DB 56 GVLMWEIFSL 65

RESULT 2  
US-08-542-363-19  
Sequence 19, Application US/08542363  
Patent No. 5770421  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,363  
FILING DATE: 12-OCT-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400001/SLE/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ. ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-542-363-19  
Query Match 100.0%; Score 52; DB 1; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GVLWEIFSL 10  
DB 56 GVLWEIFSL 65  
RESULT 3  
US-09-100-089-15  
Sequence 15, Application US/09100089  
Patent No. 6174674  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ. ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-100-089-15  
Query Match 100.0%; Score 52; DB 3; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GVLWEIFSL 10  
DB 56 GVLWEIFSL 65  
RESULT 4  
US-09-100-089-19  
Sequence 19, Application US/09100089  
Patent No. 6174674  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400002  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-100-089-19

Query Match 100.0%; Score 52; DB 3; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
|||  
DB 56 GVLLWEIFSL 65

## RESULT 5

US-09-670-827-15  
Sequence 15, Application US/09670827  
Patent No. 6451997  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-SEP-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
APPLICATION NUMBER: US 09/100,089  
FILING DATE: 19-JUN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-09-670-827-15

Query Match 100.0%; Score 52; DB 4; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
|||  
DB 56 GVLLWEIFSL 65

## RESULT 6

US-09-670-827-19  
Sequence 19, Application US/09670827  
Patent No. 6451997  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-SEP-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
APPLICATION NUMBER: US 09/100,089  
FILING DATE: 19-JUN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-670-827-19

Query Match 100.0%; Score 52; DB 4; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
|||  
DB 56 GVLLWEIFSL 65

## RESULT 7

US-09-827-949-15  
Sequence 15, Application US/09827949  
Patent No. 6696548  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof

```
FILE REFERENCE: 0656.0400004
CURRENT APPLICATION NUMBER: US/09/827,949
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: US 09/670,827
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 09/100,089
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: US 08/542,363
PRIOR FILING DATE: 1995-10-12
PRIOR APPLICATION NUMBER: US 08/160,861
PRIOR FILING DATE: 1993-12-03
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.0
SEQ ID NO 15
LENGTH: 116
TYPE: PRT
ORGANISM: Homo sapiens
US-09-827-949-15

Query Match
100.0%; Score 52; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GVLWEIFSL 10
DB 56 GVLWEIFSL 65

RESULT 8
US-09-827-949-19
Sequence 19, Application US/09827949
Patent No. 6696548
GENERAL INFORMATION:
APPLICANT: Morris, Stephan W.
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof
FILE REFERENCE: 0656.0400004
CURRENT APPLICATION NUMBER: US/09/827,949
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: US 09/670,827
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 09/100,089
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: US 08/542,363
PRIOR FILING DATE: 1995-10-12
PRIOR APPLICATION NUMBER: US 08/160,861
PRIOR FILING DATE: 1993-12-03
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.0
SEQ ID NO 19
LENGTH: 116
TYPE: PRT
ORGANISM: Homo sapiens
US-09-827-949-19

Query Match
100.0%; Score 52; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GVLWEIFSL 10
DB 56 GVLWEIFSL 65

RESULT 9
US-08-701-191A-28
Sequence 28, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schllessinger,
APPLICANT: and Stevan R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
```

```
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 290 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-28

Query Match
100.0%; Score 52; DB 2; Length 290;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GVLWEIFSL 10
DB 214 GVLWEIFSL 223

RESULT 10
US-09-664-526-28
Sequence 28, Application US/09664526
Patent No. 6682821
GENERAL INFORMATION:
APPLICANT: MOHAMMADI, MOOSA
APPLICANT: SCHLESSINGER, JOSEPH
APPLICANT: HUBBARD, STEVAN R.
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
FILE REFERENCE: 038602/0847
CURRENT APPLICATION NUMBER: US/09/664,526
CURRENT FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/188,809
PRIOR FILING DATE: 1998-11-09
PRIOR APPLICATION NUMBER: 08/701,191
PRIOR FILING DATE: 1996-08-21
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 290
TYPE: PRT
ORGANISM: Homo sapiens
US-09-664-526-28
```

Query Match 100.0%; Score 52; DB 4; Length 290;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 214 GVLWEIFSL 223

## RESULT 11

US-08-701-191A-17  
Sequence 17, Application US/08701191A  
Patent No. 5942428  
GENERAL INFORMATION:  
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,  
APPLICANT: and Stevan R. Hubbard  
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN  
TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: Storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq for Windows 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/701,191A  
FILING DATE: August 21, 1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 227/0988  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 294 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-701-191A-17

Query Match 100.0%; Score 52; DB 2; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 217 GVLWEIFSL 226

RESULT 12  
US-09-664-526-17  
Sequence 17, Application US/09664526  
Patent No. 6682321  
GENERAL INFORMATION:  
APPLICANT: MOHAMMADI, MOOSA  
APPLICANT: SCHLESSINGER, JOSEPH

APPLICANT: HUBBARD, STEVAN R.  
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN  
TITLE OF INVENTION: RECEPTOR TYROSINE KINASES  
FILE REFERENCE: 038602/0847  
CURRENT APPLICATION NUMBER: US/09/664,526  
CURRENT FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: 09/188,809  
PRIOR FILING DATE: 1998-11-09  
PRIOR APPLICATION NUMBER: 08/701,191  
PRIOR FILING DATE: 1996-08-21  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 17  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-664-526-17

Query Match 100.0%; Score 52; DB 4; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 217 GVLWEIFSL 226

RESULT 13  
US-09-390-326-5  
Sequence 5, Application US/09390326  
Patent No. 6316603  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPYAK-RUSSEL, ANNA  
APPLICANT: GERRING, MICHAEL R.  
APPLICANT: MROCKOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KAZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
TITLE OF INVENTION: METHODS OF USE  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/390,326  
CURRENT FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-390-326-5

Query Match 100.0%; Score 52; DB 3; Length 317;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 238 GVLWEIFSL 247

RESULT 14  
US-09-939-833-5  
Sequence 5, Application US/09939833  
Patent No. 6753416  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS

Job time : 25.1262 secs

APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPczyk-RUSSEL, ANNA  
APPLICANT: GERING, MICHAEL R.  
APPLICANT: MROCKOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/939,833  
CURRENT FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: US/09/390,326  
PRIOR FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-939-833-5

Query Match 100.0%; Score 52; DB 4; Length 317;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
|||  
Db 238 GVLLWEIFSL 247

RESULT 15  
US-09-506-906-5  
Sequence 5, Application US/09506906  
Patent No. 6784285  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPczyk-RUSSEL, ANNA  
APPLICANT: GERING, MICHAEL R.  
APPLICANT: MROCKOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/506,906  
CURRENT FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: US/09/390,326  
PRIOR FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-506-906-5

Query Match 100.0%; Score 52; DB 4; Length 317;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
|||  
Db 238 GVLLWEIFSL 247

Search completed: November 30, 2004, 07:19:15



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12; Search time 20 Seconds

(without alignments)  
48.108 Million cell updates/sec

Title: US-10-008-377A-6

Sequence: 1 ALPIEYGPLV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR79:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	680	4	138491
2	49	94.2	1621	2	T30200
3	41	78.8	116	2	AG2816
4	41	78.8	116	2	H97594
5	38	73.1	163	2	D82167
6	38	73.1	401	2	H93179
7	37	71.2	201	2	SS1424
8	37	71.2	1083	2	H86921
9	36	69.2	203	2	SS9279
10	36	69.2	233	2	B83544
11	36	69.2	502	2	H86264
12	36	69.2	502	2	TS2173
13	36	69.2	1078	2	PC4198
14	35	67.3	76	2	I45524
15	35	67.3	216	2	F69166
16	35	67.3	235	2	F69166
17	35	67.3	274	2	AS4951
18	35	67.3	344	2	AE6442
19	35	67.3	402	2	T48729
20	35	67.3	507	2	T08337
21	35	67.3	526	2	T04566
22	35	67.3	814	2	T00740
23	35	67.3	897	2	S76958
24	35	67.3	1207	2	H94475
25	35	67.3	7576	2	T17428
26	34	65.4	292	2	B72659
27	34	65.4	57	1	T18PVA
28	34	65.4	61	2	AF2219
29	34	65.4	77	2	F88165

30	34	65.4	123	2	AC3470	hypothetical prote
31	34	65.4	127	2	F88801	protein CO492.4 [i
32	34	65.4	127	2	T21640	hypothetical prote
33	34	65.4	127	2	T24885	hypothetical prote
34	34	65.4	127	2	F88146	protein C34F11.6 [i
35	34	65.4	127	2	A88139	protein ZK546.6 [i
36	34	65.4	127	2	C88164	protein K05F1.7 [i
37	34	65.4	127	2	H88146	protein C34F11.4 [i
38	34	65.4	127	2	G88686	protein msp-19 [im
39	34	65.4	127	2	A89165	protein ZK1248.6 [i
40	34	65.4	127	2	B88689	protein msp-65 [im
41	34	65.4	127	2	F88138	protein msp-31 [im
42	34	65.4	127	2	G88145	protein F58A6.8 [i
43	34	65.4	127	2	D88164	protein msp-142 [i
44	34	65.4	127	2	H88688	protein msp-59 [im
45	34	65.4	127	2	C88689	protein msp-51 [im

#### ALIGNMENTS

##### RESULT 1

138491  
nucleophosmin/ anaplastic lymphoma kinase mutant fusion protein - human

C/Species: Homo sapiens (man)  
C/Date: 15-Feb-1996 #sequence\_revision 15-Feb-1996 #text\_change 20-Apr-2000

C/Accession: 138491  
R/Morris, S.W.; Kirstein, M.N.; Valentine, M.B.; Dittmer, K.G.; Shapiro, D.N.; Saltman,

Science 263, 1281-1284, 1994  
A/Title: Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin

A/Reference number: A53182; MIMD:94167588; PMID:8122112  
A/Accession: 138491

A/Molecule type: mRNA  
A/Residues: 1-680 <MOR>

A/Cross-references: EMBL:U04946; NID:G609341; PIDN:AA58698.1; PID:G609342  
C/Comment: This sequence is the chimeric product of a translocation mutation.

C/Genetics:  
A/Gene: NPM/ALK  
A/Map position: 5/2p23-2p23

C/Keywords: fusion protein

Query Match	Score	DB	Length
100.0%;	52;	4;	680;
Best Local Similarity	100.0%;	Pred. No. 0.058;	
Matches	10;	Conservative	0;
		Mismatches	0;
		Indels	0;
		Gaps	0;

##### RESULT 2

T30200  
protein-tyrosine kinase (EC 2.7.1.112) alk - mouse

N/Alternate names: anaplastic lymphoma kinase  
A/Species: Mus musculus (house mouse)

C/Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T30200

R/Inahara, T.; Fujimoto, J.; Wen, D.; Cupples, R.; Bucay, N.; Arakawa, T.; Mori, S.; Ra

Oncogene 14, 439-449, 1997  
A/Title: Molecular characterization of ALK, a receptor tyrosine kinase expressed specif

A/Reference number: Z20774; MIMD:97178863; PMID:9033841  
A/Accession: T30200

A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA

A/Residues: 1-1621 <TMA>  
A/Cross-references: UNIPROT:P97793; EMBL:DB3002; NID:q186406; PIDN:BA11673.1; PID:918

A/Experimental source: brain and testis  
C/Genetics:

A/Gene: alk  
C/Function:

A/Description: may play an important role in development of the brain  
C/Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

Query Match 94.2%; Score 49; DB 2; Length 1621;  
 Best Local Similarity 90.0%; Pred. No. 0.58;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPL 10  
 |||||  
 DB 1400 ALPIEYGPLV 1409

## RESULT 3

AG2816 conserved hypothetical protein Atu1953 [imported] - Agrobacterium tumefaciens (strain C58)

C|Species: Agrobacterium tumefaciens  
 C|Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
 C|Accession: AG2816  
 R|Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
 erage, G.; Gillet, W.; Grant, C.; Genthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A|Authors: Yoo, H.; Tao, Y.; Biddle, P.; Ung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, R.W.  
 A|Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A|Reference number: AB2577; MIMD:21608550; PMID:11743193  
 A|Accession: AG2816  
 A|Status: preliminary  
 A|Molecule type: DNA  
 A|Residues: 1-116 <KUR>  
 A|Cross-references: UNIPROT:Q8UE11; GB:AE008688; PIRN:AA142949.1; PIR:G17740407; GSPDB:G  
 A|Experimental source: strain C58 (Dupont)  
 C|Genetics:  
 A|Gene: Atu1953  
 A|Map position: circular chromosome

Query Match 78.8%; Score 41; DB 2; Length 116;  
 Best Local Similarity 87.5%; Pred. No. 1.1;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPIEYGPL 9  
 |||||  
 DB 84 LPPIEYGPLV 91

## RESULT 4

H97594 hypothetical protein AGR\_C\_3564 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C|Species: Agrobacterium tumefaciens  
 C|Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
 C|Accession: H97594  
 R|Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Mollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A|Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A|Reference number: A97359; MIMD:21608551; PMID:11743194  
 A|Accession: H97594  
 A|Status: preliminary  
 A|Molecule type: DNA  
 A|Residues: 1-116 <KUR>  
 A|Cross-references: UNIPROT:Q8UE11; GB:AE007869; PIRN:AAK67713.1; PIR:G15157076; GSPDB:G  
 C|Genetics:  
 A|Gene: AGR\_C\_3564  
 A|Map position: circular chromosome

Query Match 78.8%; Score 41; DB 2; Length 116;  
 Best Local Similarity 87.5%; Pred. No. 1.1;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPIEYGPL 9  
 |||||  
 DB 84 LPPIEYGPLV 91

## RESULT 5

DB2167

conserved hypothetical protein VC1703 [imported] - Vibrio cholerae (strain N1696; serogr

C|Species: Vibrio cholerae  
 C|Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C|Accession: DB2167  
 R|Heidelberg, U.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoti, I.; Sellers, F  
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A|Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A|Reference number: AB2035; MIMD:20406833; PMID:10952301  
 A|Accession: DB2167  
 A|Status: preliminary  
 A|Molecule type: DNA  
 A|Residues: 1-163 <HRI>  
 A|Cross-references: UNIPROT:Q3KRD9; GB:AE004248; GB:AE003852; NID:g9656216; PIRN:AAE948  
 A|Experimental source: serogroup O1, strain N16961; biotype El Tor  
 C|Genetics:  
 A|Gene: VC1703  
 A|Map position: 1

Query Match 73.1%; Score 38; DB 2; Length 163;  
 Best Local Similarity 75.0%; Pred. No. 6.3;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPIEYGPL 9  
 |||||  
 DB 76 LPPIEYGPL 83

## RESULT 6

P83179 hypothetical protein PA3734 [imported] - Pseudomonas aeruginosa (strain PA01)

C|Species: Pseudomonas aeruginosa  
 C|Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C|Accession: P83179  
 R|Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.T.; B  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Tachib, K.; Lim,  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A|Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
 A|Reference number: AB2950; MIMD:20437337; PMID:10964043  
 A|Accession: P83179  
 A|Status: preliminary  
 A|Molecule type: DNA  
 A|Residues: 1-401 <STO>  
 A|Cross-references: UNIPROT:Q9HXQ9; GB:AE004732; GB:AE004091; NID:g9949895; PIRN:AA6071  
 A|Experimental source: strain PA01  
 C|Genetics:  
 A|Gene: PA3734

Query Match 73.1%; Score 38; DB 2; Length 401;  
 Best Local Similarity 88.9%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALPIEYGPL 9  
 |||||  
 DB 278 ALPIEYGPL 286

## RESULT 7

S51424 hypothetical protein YLR179c - yeast (Saccharomyces cerevisiae)

N|Alternate names: hypothetical protein L9470.20  
 C|Species: Saccharomyces cerevisiae  
 C|Date: 23-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
 C|Accession: S51424  
 R|Mohlmann, P.  
 submitted to the EMBL Data Library, November 1994  
 A|Description: The sequence of S. cerevisiae cosmid 9470.  
 A|Reference number: S5144  
 A|Accession: S51424  
 A|Molecule type: DNA  
 A|Residues: 1-201 <WOH>

A;Cross-references: UNIPROT:Q66252; EMBL:U117246; NID:G577192; FIDN:AA667472.1; PID:G57772  
A;Gene: MRPS.YLR179C  
A;Cross-references: SGD:S0004169  
A;Map position: 12R

Query March	71.2%;	Score 37;	DB 2;	Length 201;
Best Local Similarity	71.4%;	Pred. No. 12;		
Matches	5;	Conservative	2;	Mismatches
			0;	Indels
				Gaps
QY	2	LP:EVGP	8	
Db	101	IP:EVGP	107	

```

RESULT 8
H86921
probable arabinosyl transferase [imported] - Mycobacterium leprae
C|Species: Mycobacterium leprae
C|Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C|Accession: AF000001

```

R; Cole, S.T.; Elismatier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holt,  
C; Accession: hb6721  
R; Davies, R.M.; Devlin, K.; Dutchoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
Eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A; Authors: Butler, S.; Seeger, K.; Simon, S.; Symmonds, M.; Skelton, J.; Squares, R.; Se  
A; Title: Massive gene decay in the leprosy bacillus.  
A; Accession: H86921  
A; Reference number: A86909; MCID:21128732; PMID:11234002  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-1083 <STO>  
A; Cross-references: UNIPROT:Q9CDA9; GB:AL450360; NID:g13092485; P1DN:CAC29612.1; GSPDB:  
C; Genetics  
A; Gene: embB  
A; Superfamily: probable arabinosyl transferase

Query Match	71.2%	Score 37	DB 2	Length 1083
Best Local Similarity	77.8%	Pred. No. 79		
Matches	7	Conservative	0	Mismatches 2; Indels 0; Gaps 0;

```
QY      1 ALPIEYGPL 9
        ||| |||
Db      753 ALPSNYGPL 761
```

RESULT 9  
S59279  
hypothetical protein YDR041w - yeast (Saccharomyces cerevisiae)  
N|Alternate names: hypothetical protein YD8688.03  
C|Species: Saccharomyces cerevisiae  
C|Date: 29-Nov-1995 #sequence\_revision 01-Mar-1996 #text\_change 09-Jul-2004  
Accession: S59279

R:Murphy, U.; Harris, D.  
Submitted to the EMBL Data Library, September 1995  
A:Reference number: S59277  
A:Accession: S59279  
A:Molecule type: DNA  
A:Residues: 1-203 <NUR>  
A:Cross-references: UNIPROT:Q03201; EMBL:Z54075; NID:G976128; PID:G976131; GSPDB:GN00004  
C:Genetics:  
A:Gene: SGD:RSM10; MIPS:YDR041w  
A:Cross-references: SGD:S0002448  
A:Map position: 4R  
C:Superfamily: ribosomal protein S10/S20

Query Match	69.2%;	Score 36;	DB 2;	Length 203;
Best Local Similarity	77.8%;	Pred. No. 20;		
Matches	7;	Conservative	1;	Mismatches 1; Indels 0; Gaps 0;

```

QY      2 LPIEYGPLV 10
         |||:||||
Db      34 LPIKYGDLV 42

```

RESULT 10  
B83544

probable halosulfate dehydrogenase PA0810 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 03-Jul-2004

C:Accession: B83594  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warener, P.; Hickey, M.J.; E  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; Lim  
:; Lory, S.; Olson, M.V.  
Nature 406: 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic path  
A:Reference number: A82950; PMID:20437537; PMID:10584043

A|Accession: B93544  
A|Status: preliminary  
A|Molecule type: DNA  
A|Residues: 123 <ro>  
A|Cross-references: UNIPROT:Q9ISC9, GB:AE004516, GB:AE004091, NID:5946699, FIDN:AA0411  
A|Experimental source: strain PA01  
C|Genetics:  
A|Gene: PA0810

Query Match	69.2%	Score 36;	DB 2;	Length 233;
Best Local Similarity	75.0%;	Pred. No. 23;		
Matches	6;	Conservative	1;	Mismatches 0;
				Gaps 0;

```

QY      1 ALPIEYGP 8
        | | | | |
Db      196 ARPLEYGP 203

```

RESULT 11  
H86264

cytochrome P450 monooxygenase F3F19.10 - *Arabidopsis thaliana*  
C1Species: *Arabidopsis thaliana* (mouse-ear cress)  
C1Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C; Accession: H8264  
R; Theologian, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
ansen, N.F.; Hughes, B.; Hulzar, L.  
Nature 408, 815-820, 2000  
A; Authors: Hunter, J.L.; Jenkins, T.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, U.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, U.S.; Mailli, R.; Marzall  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A; Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A; Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A; Reference number: A86141; M0121016719; PMID11130712

A1:Accession: H86264  
A1:Status: preliminary  
A1:Molecule type: DNA  
A1:Residues: 1-502 <10>  
A1:Cross-references: UNIPROT:O65788, GB:AE005172, NID:AA030391, PIDD:AA031061.1, GSPDB:G  
A1:Genetics:  
A1:Map position: 1  
A1:Superfamily: human cytochrome P450 CYP2D6, cytochrome P450 homology  
A1:Keywords: heme, iron, metalloprotein  
A1:444/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match	69.2%;	Score 36;	DB 2;	Length 502;
Best Local Similarity	77.8%;	Pred. No. 53;		
Matches	7;	Conservative	1;	Indels 0; Gaps 0;

```
QY      2 LPIEYGPLV 10
          | : |||||
Db      58 LSIKYGPLV-66
```

RESULT 12

152173 cytochrome P450 monooxygenase [imported] - *Arabidopsis thaliana*  
C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004

C/Accession: T52173

R/Mutnant: M.; Ward, E.; Ohta, D.

Plant Mol. Biol. 37, 39-52, 1998

A/Title: Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of cDNAs, differ

A/Reference number: 214382; PMID:98281573; PMID:9620263

A/Accession: T52173

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-502 <M1Z>

A/Cross-references: UNIPROT:O65788; EMBL:D78605; PDB:1BPA28537.1

C/Genetics:

A/Genes: CYP71B2

C/Suprafamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C/Keywords: heme; iron; metalloprotein

F/301-466/Domain: cytochrome P450 homology <P45>

F/444/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 69.2%; Score 36; DB 2; Length 502;

Best Local Similarity 77.8%; Pred. No. 53;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIRYGPLY 10

DB 58 LPIRYGPLY 66

#### RESULT 13

peptide synthetase (EC 1.14.17.-) - imperfect fungus (Metarhizium anisopliae) (fragment)

C/Species: Metarhizium anisopliae

C/Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 09-Jul-2004

C/Accession: PC4198; PC4199

R/Balley, A.M.; Kershaw, M.J.; Hunt, B.A.; Paterson, I.C.; Charnley, A.K.; Reynolds, S.E

Gene 173, 195-197, 1996

A/Title: Cloning and sequence analysis of an intron-containing domain from a peptide syn

A/Reference number: PC4198; PMID:97082966; PMID:8964498

A/Accession: PC4198

A/Molecule type: DNA

A/Residues: 1-1078 <BAI>

A/Cross-references: UNIPROT:Q01135; EMBL:X89442; NID:G1015536; PDB:G1015537

A/Accession: PC4199

A/Molecule type: protein

A/Residues: 58-67/136-146 <BIZ>

C/Genetics:

A/Genes: pes

A/Inserts: 386/3; 762/2

C/Keywords: carrier protein; oxidoreductase

F/32-499/Domain: acetate-CoA ligase homology <ACD>

F/526-594/Domain: acyl carrier protein homology <ACP3>

Query Match 69.2%; Score 36; DB 2; Length 1078;

Best Local Similarity 77.8%; Pred. No. 1,2e+02;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIRYGPLY 10

DB 729 LPIRYGPLY 737

#### RESULT 14

142524

A/ORF-M protein - vaccinia virus (strains WR and Copenhagen)

N/Alternate names: SALB protein

C/Species: vaccinia virus

A/Note: host Homo sapiens (man)

C/Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 09-Jul-2004

C/Accession: 142524; J01819

R/Johnson, G.P.

submitted to GenBank, June 1990

A/Reference number: A33172

A/Accession: 142524

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-76 <J0H>

A/Cross-references: UNIPROT:P20522

R/Smith, G.L.; Chan, Y.S.; Howard, S.T.

J. Gen. Virol. 72, 1349-1376, 1991

A/Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right ir

A/Reference number: J01767; PMID:91259063; PMID:2045793

A/Accession: J01819

A/Molecule type: DNA

A/Residues: 1-76 <SMI>

A/Cross-references: DDBJ:D11079

Query Match 67.3%; Score 35; DB 2; Length 76;

Best Local Similarity 75.0%; Pred. No. 10;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIRYGPLY 9

DB 17 LPIRYGPLY 24

#### RESULT 15

F69166

hypothetical protein MTH506 - Methanobacterium thermoautotrophicum (strain Delta H)

C/Species: Methanobacterium thermoautotrophicum

C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C/Accession: F69166

R/Smith, D.R.; Doucette-Stamm, L.A.; DeJonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.

Xi, S.; Chun, G.M.; Daniels, C.D.; Mao, J.; Rice, S.; Noelling, U.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A/Reference number: A69000; PMID:98037514; PMID:9371463

A/Accession: F69166

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-216 <MTH>

A/Cross-references: UNIPROT:O26606; GB:AE000834; GB:AE000666; NID:G2621574; PDB:AAE501

A/Experimental source: strain Delta H

C/Genetics:

A/Genes: MTH506

A/Start codon: GTG

Query Match 67.3%; Score 35; DB 2; Length 216;

Best Local Similarity 71.4%; Pred. No. 33;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRYGP 8

DB 143 VPIRYGP 149

Search completed: November 30, 2004, 07:41:24  
Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 76.4103 Seconds

(without alignments)  
75.301 Million cell updates/sec

Title: US-10-008-377A-6  
Perfect score: 1 ALPIEXPLV 10  
Sequence: 1 ALPIEXPLV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: UniProt 02: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	52	100.0	803 2 Q8TDJ5	Q8TDJ5 homo sapien
2	52	100.0	1620 1 ALK_HUMAN	Q8TDJ5 homo sapien
3	49	94.2	1621 1 ALK_MOUSE	P97793 mus musculu
4	44	84.6	130 2 Q7MM63	Q7MM63 vibrio vuln
5	44	84.6	130 2 Q8D8A5	Q8D8A5 vibrio vuln
6	42	80.8	130 2 Q8D8A2	Q8D8A2 vibrio para
7	41	78.8	116 2 Q7C171	Q7C171 agrobacteri
8	41	78.8	116 2 Q8D8E1	Q8D8E1 agrobacteri
9	41	78.8	339 2 Q8XXA2	Q8XXA2 anopheles s
10	40	76.9	529 2 Q7PV84	Q7PV84 anopheles g
11	40	76.9	1721 2 Q9UNR3	Q9UNR3 pseudomonas
12	3277	76.9	2 07NN14	07NN14 streptomyce
13	39	75.0	2 Q8FN67	Q8FN67 glycinibact
14	38	73.1	144 2 Q8SD08	Q8SD08 pseudomonas
15	38	73.1	163 2 Q8KRD9	Q8KRD9 vibrio chol
16	38	73.1	401 2 Q8HXQ8	Q8HXQ8 pseudomonas
17	37	71.2	61 2 Q8ILR2	Q8ILR2 plasmodium
18	37	71.2	201 1 YL79_YEAST	YL79_YEAST
19	37	71.2	239 2 Q89M29	Q89M29 bradyrhizob
20	37	71.2	306 2 Q53929	Q53929 streptomyce
21	37	71.2	2 08CUC3	08CUC3 streptomyce
22	37	71.2	673 2 Q7OEW1	Q7OEW1 glycine max
23	37	71.2	2 07OEW1	07OEW1 glycine max
24	37	71.2	1083 1 EMBR_MYCLE	EMBR_MYCLE
25	37	71.2	7525 2 Q8KIE0	Q8KIE0 mycobacteri
26	36	69.2	167 2 Q8GCA1	Q8GCA1 streptomyce
27	36	69.2	203 2 Q03201	Q03201 streptomyce
28	36	69.2	203 2 AASS5974	AASS5974 saccharomyc
29	36	69.2	233 2 Q915C9	Q915C9 pseudomonas
30	36	69.2	298 2 Q7X295	Q7X295 streptomyce
31	36	69.2	322 2 Q7Q062	Q7Q062 anopheles g

32	36	69.2	328 2 Q7M382	Q7M382 bordetella
33	36	69.2	328 2 Q7MEK1	Q7MEK1 bordetella
34	36	69.2	433 2 Q8D9F0	Q8D9F0 xanthomonas
35	36	69.2	498 2 Q76M94	Q76M94 asparagus
36	36	69.2	498 2 BAD06417	BAD06417 asparagus
37	36	69.2	502 1 C722_ARATH	C722_ARATH
38	36	69.2	503 2 Q75749	Q75749 muscardi a
39	36	69.2	503 2 BAD16679	BAD16679 muscardi a
40	36	69.2	503 2 BAD16680	BAD16680 muscardi a
41	36	69.2	619 2 Q9X4G3	Q9X4G3 aquilex pyr
42	36	69.2	963 2 Q8KWC7	Q8KWC7 rubegetia sp
43	36	69.2	1670 1 CLOT_PENMO	CLOT_PENMO
44	36	69.2	5157 2 Q01135	Q01135 penaeus mon
45	35.5	68.3	152 2 Q98NC2	Q98NC2 rhabdium 1

## ALIGNMENTS

RESULT 1  
Q8TDJ5 PRELIMINARY: PRT: 803 AA.  
AC Q8TDJ5  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE TRK-fused gene/anaplastic large cell lymphoma kinase extra long form.  
GN Name=TRF/ALK fusion;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21940189; PubMed=11943732;  
RA Hernandez L., Bea S., Bellosillo B., Pinyol M., Falini B., Ott G., Pulford K., Rosenwald A., Morris S.W., Fernandez A., Santos E., Campo E.,  
RA "Diversity of genomic breakpoints in TRF-ALK translocations in anaplastic large cell lymphomas: identification of a new TRF-ALK(XL) chimeric gene with transforming activity",  
Am. J. Pathol. 160:1487-1494(2002).  
DR EMBL: AF190893; AAM17922.1; -  
DR HSSP: Q62838; 1LUF.  
DR GO: GO:0005524; F:ATP binding; IEA.  
DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO: GO:0016740; F:transferase activity; IEA.  
DR GO: GO:0006468; F:protein amino acid phosphorylation; IEA.  
DR InterPro: IPR011009; Kinase-like.  
DR InterPro: IPR000270; GPR\_PBI.  
DR InterPro: IPR000719; Prot\_Kinase.  
DR InterPro: IPR002011; ReceptTyknsII.  
DR InterPro: IPR001245; Tyr\_Pkinase.  
DR InterPro: IPR008265; Tyr\_Pkinase\_AS.  
DR Pfam: PF00564; PBI.1.  
DR Pfam: PF00069; Ekinase.1.  
DR PRINTS: PR00109; TRKINASE.  
DR PRODOM: PD000001; Prot\_Kinase.1.  
DR SMART: SM00666; PBI.1.  
DR SMART: SM00219; TykC.1.  
DR PROSITE: PS00107; PROTEIN KINASE ATP; UNKNOWN\_1.  
DR PROSITE: PS00011; PROTEIN KINASE DOM; 1.  
DR PROSITE: PS00109; PROTEIN KINASE TYR; 1.  
DR PROSITE: PS00239; RECEPTOR TYR\_KIN\_II; UNKNOWN\_1.  
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
SQ SEQUENCE 803 AA; 86670 MW; 890826d95b695f CRC64;  
Query Match 100.0%; Score 52; DB 2; Length 803;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ALPIEXPLV 10

DB 579 ALPIEYGBLY 568

RESULT 2  
 ALK\_HUMAN STANDARD; PRT; 1620 AA.  
 AC Q9UM73; Q9YAK6;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE ALK tyrosine kinase receptor precursor (EC 2.7.1.112) (Amplastic  
 DE lymphoma kinase) (CD246 antigen).  
 GN Name:ALK;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND  
 RP GLYCOSYLATION.  
 RX MEDLINE=97316773; PubMed=9174053; DOI=10.1038/51.000.1201062;  
 RA Morris S.W., Naeve C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,  
 RT Witte D.P.;  
 RT "ALK, the chromosome 2 gene locus altered by the t(2;5) in non-  
 RT Hodgkin's lymphoma, encodes a novel neutral receptor tyrosine kinase  
 RT that is highly related to leukocyte tyrosine kinase (LTK)."  
 RL Oncogene 14:2175-2188(1997).  
 [2]  
 RP ERRATUM.  
 RA Morris S.W., Naeve C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,  
 RA Witte D.P.;  
 RL Oncogene 15:2883-2883(1997).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97178663; PubMed=9053841;  
 RA Iwahara T., Fujimoto J., Wen D., Cupples R., Bucay N., Arakawa T.,  
 RA Mori S., Ratakin B., Yamamoto T.;  
 RT "Molecular characterization of ALK, a receptor tyrosine kinase  
 RT expressed specifically in the nervous system."  
 RL Oncogene 14:439-449(1997).  
 [4]  
 RP PARTIAL SEQUENCE FROM N.A., AND CHROMOSOMAL TRANSLOCATION.  
 RX MEDLINE=94167588; PubMed=8122112;  
 RA Morris S.W., Kirstein M.N., Valentine M.B., Dittmer K.G.,  
 RA Shapiro D.N., Saltman D.L., Look A.T.;  
 RT "Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in  
 RT non-Hodgkin's lymphoma."  
 RL Science 263:1281-1284(1994).  
 CC -1- FUNCTION: Orphan receptor with a tyrosine-protein kinase activity.  
 CC Appears to play an important role in the normal development and  
 CC function of the nervous system.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Expressed in brain and CNS. Also expressed in  
 CC the small intestine and testis, but not in normal lymphoid cells.  
 CC -1- PTM: N-glycosylated.  
 CC -1- DISEASE: A form of non-Hodgkin's lymphoma is characterized by a  
 CC chromosomal translocation t(2;5)(p23;q35) that involves NPM and  
 CC ALK.  
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin  
 CC receptor subfamily.  
 CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.  
 CC -1- SIMILARITY: Contains 2 MAM domains.  
 CC -1- DATABASE: NMRAtlas Genet. Cyrogenet. OncoJ. Haematol.;  
 CC WWW="http://www.inobiochem.fr/services/chromocancer/Genes/ALK.html".  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL, U62540; AAB71619.1; -  
 DR EMBL, U66559; AAC51104.1; -  
 DR HSSP, Q62838; ILOF.  
 DR Genew; HGNC:427; ALK.  
 DR MIM; 105530; -  
 DR GO; GO:0016021; C: integral to membrane; NAS.  
 DR GO; GO:0007399; P: neurogenesis; NAS.  
 DR InterPro; IPR011009; Kinase like.  
 DR InterPro; IPR002172; LDL receptor\_A.  
 DR InterPro; IPR000998; LAM receptor\_A.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR002011; ReceptTykinst1.  
 DR InterPro; IPR01245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00629; MAM; 1.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00192; Ldla; 1.  
 DR SMART; SM00239; Tyrc; 1.  
 DR PROSITE; PS01209; LDLR\_1; FALSE NEG.  
 DR PROSITE; PS00068; LDLR\_2; FALSE NEG.  
 DR PROSITE; PS00740; MAM\_1; FALSE NEG.  
 DR PROSITE; PS00060; MAM\_2; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00039; RECEPTOR\_TYR\_KIN\_II; 1.  
 DR ATP-binding; Chromosomal translocation; Glycoprotein; Phosphorylation;  
 DR Proto-oncogene; Receptor; Repeat; Signal; Transferase; Transmembrane;  
 DR Tyrosine-protein kinase.  
 KW SIGNAL  
 FT CHAIN 1 18  
 FT 19 1620  
 FT DOMAIN 19 1038  
 FT TRANSMEM 1039 1059  
 FT DOMAIN 1060 1620  
 FT DOMAIN 264 427  
 FT DOMAIN 437 473  
 FT DOMAIN 478 636  
 FT DOMAIN 1116 1392  
 FT DOMAIN 816 940  
 FT NP\_BIND 1122 1130  
 FT BINDING 1150 1150  
 FT ACT\_SITE 1249 1249  
 FT MOD\_RES 1282 1282  
 FT CARBOHYD 169 169  
 FT CARBOHYD 244 244  
 FT CARBOHYD 285 285  
 FT CARBOHYD 324 324  
 FT CARBOHYD 411 411  
 FT CARBOHYD 424 424  
 FT CARBOHYD 445 445  
 FT CARBOHYD 563 563  
 FT CARBOHYD 571 571  
 FT CARBOHYD 627 627  
 FT CARBOHYD 709 709  
 FT CARBOHYD 808 808  
 FT CARBOHYD 863 863  
 FT CARBOHYD 864 864  
 FT CARBOHYD 886 886  
 FT CARBOHYD 986 986  
 FT CONFLICT 36 36  
 FT CONFLICT 1491 1491  
 FT CONFLICT 1529 1529  
 SQ SEQUENCE 1620 AA; 176417 MW; A62604B24296151E CRC64;  
 Query Match 100.0%; Score 52; DB 1; Length 1620;  
 Best Local Similarity 100.0%; Pred. No. 0.87;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 ALPTEVGPV 10  
DB 1396 ALPTEVGPV 1405

RESULT 3  
ID ALK\_MOUSE STANDARD; PRT; 1621 AA.

AC P97793;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE ALK tyrosine kinase receptor precursor (EC 2.7.1.112) (Anaplastic  
DE lymphoma kinase).  
GN Name=Alk;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain, and Testis;  
RX MEDLINE=9717863; PubMed=9053841;  
RA Iwahara T., Fujimoto Y., Wen D., Cupples R., Bucay N., Arakawa T.,  
RA Mori S., Ratzkin B., Yamamoto T.;  
RT "Molecular characterization of Alk, a receptor tyrosine kinase  
RT expressed specifically in the nervous system."  
RL Oncogene 14:439-449(1997).  
CC -1- FUNCTION: Orphan receptor with a tyrosine-protein kinase activity.  
CC Appears to play an important role in the normal development and  
CC function of the nervous system.  
CC Catalytic Activity: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).  
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin  
CC receptor subfamily.  
CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.  
CC -1- SIMILARITY: Contains 2 MAM domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D83002; BA011673.1; -;  
DR PIR; T30200; T30200.  
DR HSSP; P08063; L1QH.  
DR WGI; MGI:103305; Alk.  
DR GO; GO:0005515; F:Protein binding; IPT.  
DR InterPro; IPR011008; Kinase\_1ike.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR000998; MAM.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR002011; ReceptTykinst1.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00628; MAM; 1.  
DR Pfam; PF00069; Kinase\_1.  
DR PRINTS; PR0109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00192; Ldla; 1.  
DR SMART; SM00219; TYRK; 1.  
DR PROSITE; PS01209; LDLRA\_1; FALSE NEG.  
DR PROSITE; PS00668; LDLRA\_2; FALSE NEG.  
DR PROSITE; PS00740; MAM\_1; FALSE NEG.  
DR PROSITE; PS00660; MAM\_2; 2.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS0011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.

DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
KW ATP-binding; Glycoprotein; Phosphorylation; Receptor; Repeat; Signal;  
KW Transferase; Transmembrane; Tyrosine-protein kinase.  
FT SIGNAL 1  
FT CHAIN 1  
FT DOMAIN 19 1621  
FT TRANSMEM 19 1042  
FT DOMAIN 1043 1063  
FT DOMAIN 1064 1621  
FT DOMAIN 268 431  
FT DOMAIN 441 477  
FT DOMAIN 482 640  
FT DOMAIN 1120 1396  
FT DOMAIN 944  
FT NP\_BIND 1126 1134  
FT BINDING 1154 1154  
FT ACT\_SITE 1253 1253  
FT MOD\_RES 1286 1286  
FT CARBOHYD 174 174  
FT CARBOHYD 248 248  
FT CARBOHYD 289 289  
FT CARBOHYD 328 328  
FT CARBOHYD 415 415  
FT CARBOHYD 428 428  
FT CARBOHYD 449 449  
FT CARBOHYD 567 567  
FT CARBOHYD 575 575  
FT CARBOHYD 631 631  
FT CARBOHYD 673 673  
FT CARBOHYD 713 713  
FT CARBOHYD 812 812  
FT CARBOHYD 868 868  
FT CARBOHYD 890 890  
FT CARBOHYD 990 990  
SQ SEQUENCE 1621 AA; 174919 MW; 16E25BE21AADE22 CRC64;

Query Match  
Best Local Similarity 90.0%; Score 49; DB 1; Length 1621;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 1 ALPTEVGPV 10  
DB 1400 ALPTEVGPV 1409

RESULT 4  
ID Q7MM63 PRELIMINARY; PRT; 130 AA.

AC Q7MM63;  
DT 01-MAR-2004 (TRENBERG; 26, Created)  
DT 01-MAR-2004 (TRENBERG; 26, Last sequence update)  
DT 01-MAR-2004 (TRENBERG; 26, Last annotation update)  
DE Hypothetical protein V1210.  
GN Name=V1210;  
OS Vibrio vulnificus (strain V1016).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrrio.  
OX NCBI\_TaxID=196600;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=1456365;  
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,  
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li D.-C., Su T.-L.,  
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;  
RT "Comparative genome analysis of Vibrio vulnificus, a marine  
RT pathogen."  
RL Genome Res. 13:2577-2587(2003).  
DR EMBL; AP005334; BAC93974.1; -;  
KM Hypothetical protein.  
SQ SEQUENCE 130 AA; 14171 MW; 6C1BC5D7EC734063 CRC64;

Query Match  
Best Local Similarity 94.6%; Score 44; DB 2; Length 130;  
Matches 100.0%; Pred. No. 2.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEXGPL 9  
DB 43 LPIEXGPL 50

## RESULT 5

Q8D8A5 PRELIMINARY; PRT; 130 AA.  
AC Q8D8A5;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocustNames=V13076;  
OS *Vibrio vulnificus*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; *Vibrio*.  
OX NCBI\_TaxID=672;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CMP6;  
RA Rhee U.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
RA Choe H.E.;  
RT "Complete genome sequence of *Vibrio vulnificus* CMP6.";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL, AE016807; AAC11400.1;  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 130 AA; 14803 MW; 6C1B5DDEC9F1C8 CRC64;

Query Match 84.6%; Score 44; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEXGPL 9  
DB 43 LPIEXGPL 50

## RESULT 6

Q87NA2 PRELIMINARY; PRT; 130 AA.  
AC Q87NA2;  
DT 01-JUN-2003 (TREMBlrel. 24, Created)  
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Hypothetical protein VPI973.  
GN OrderedLocustNames=VPI973;  
OS *Vibrio parahaemolyticus*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; *Vibrio*.  
OX NCBI\_TaxID=670;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IMD 2210633 / Serotype O3:K6;  
RX MEDLINE=22508454; PubMed=12620739;  
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
RA Iijima Y., Nishimura M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;  
RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism  
distinct from that of *V. cholerae*.";  
RL Lancet 361:743-749 (2003).  
DR EMBL, AP005079; BAC60236.1;  
DR InterPro: IPR01047; Quin a1c DH like.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 130 AA; 14412 MW; 7E3F1A9DE506AC CRC64;

Query Match 80.8%; Score 42; DB 2; Length 130;  
Best Local Similarity 87.5%; Pred. No. 5.7;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEXGPL 9

DB 42 LPIEXGPL 49

## RESULT 7

Q7CY71 PRELIMINARY; PRT; 116 AA.  
AC Q7CY71;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE AGR C 3564D.  
GN OrderedLocustNames=AGR C 3564;  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
OX NCBI\_TaxID=176299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Cereon;  
RX MEDLINE=21608551; PubMed=11743194;  
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
RA Curcio B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
RA Hummel K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,  
RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,  
RA Flanagan C., Crowell C., Garsen J., Lomo C., Seer C., Strub G.,  
RA Cielo C., Slater S.;  
RT "Genome sequence of the plant pathogen and biotechnology agent  
Agrobacterium tumefaciens C58.";  
RL Science 294:2323-2328 (2001).  
DR EMBL, AE008113; AAK8713.1;  
DR InterPro: IPR004360; Gly\_b1ox.  
DR Pfam: PF00903; Glyoxalase\_1.  
SQ SEQUENCE 116 AA; 12585 MW; 9FA61753836C7655 CRC64;

Query Match 78.8%; Score 41; DB 2; Length 116;  
Best Local Similarity 87.5%; Pred. No. 7.9;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEXGPL 9  
DB 84 LPIEXGPL 91

## RESULT 8

Q8UE11 PRELIMINARY; PRT; 116 AA.  
AC Q8UE11;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Hypothetical protein Adu1953.  
GN OrderedLocustNames=Adu1953;  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
OX NCBI\_TaxID=176299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Dupont;  
RX MEDLINE=21608550; PubMed=11743193;  
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.F.,  
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,  
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
RA Chapman P., Clendenning J., Deatherage G., Gilliet W., Grant C.,  
RA Kutyavlin T., Levy R., Li M.-J., McClelland E., Palmeri A.,  
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,  
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
RA Chunley F., Tingey S.V., Tomo J.-F., Gordon M.P., Olson M.V.,  
RA Nester E.W.;  
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
C58.";



```

RL Science 294:2317-2323(2001).
DR EMBL; A6009148; AAL42949.1; -.
DR PIR; AG2816; AG2816.
DR PIR; H97594; H97594.
DR InterPro; IPR004360; Gly_bleo_dtox.
DR Pfam; PF00903; Glyoxalase.1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 116 AA; 12585 MW; 9FA6175386C7655 CRC64;

Query Match
Best Local Similarity 78.8%; Score 41; DB 2; Length 116;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEYGPL 9
DB 84 LPIEYGPLV 91

RESULT 9
Q8XXA2 PRELIMINARY; PRT; 339 AA.
AC Q8XXA2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein RSC2214
GN Name=RSC2214; OrderedLocNames=RSC2214;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM1000;
RX Medline=21681879; Pubmed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Aizac M., Billault A., Brocher P., Camus J.C., Catolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cumac S., Demange N.,
RA Gaspin C., Lavie W., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigler P., Thebaud P., Whalen M., Winkler P., Levy M.,
RA Weissbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646068; CADI5921.1; -.
DR InterPro; IPR001279; Blackmase-like.
DR Pfam; PF00753; Lactamase_B; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 339 AA; 36167 MW; 85684DE2A5DCDD26 CRC64;

Query Match
Best Local Similarity 78.8%; Score 41; DB 2; Length 339;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEYGPL 9
DB 326 VPIEYGPL 333

RESULT 10
Q7PVB4 PRELIMINARY; PRT; 529 AA.
AC Q7PVB4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000016583 (Fragment).
GN Name=ENSANGP0000016583;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anophelinae.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008985; EAA00286.2; -.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR01547; Sulph_transpt.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR PROSITE; PS50801; STAS; 1.
FT NON TER 1
SQ SEQUENCE 529 AA; 57283 MW; 9A037334418341ED CRC64;

Query Match
Best Local Similarity 76.9%; Score 40; DB 2; Length 529;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10
DB 382 ALPIEYGPLV 391

RESULT 11
Q9UNR3 PRELIMINARY; PRT; 1721 AA.
AC Q9UNR3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Clotting protein precursor.
OS Pacifastacus leniusculus (Signal crayfish).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Decapoda; Decapoda; Pleocyemata; Astacidea;
OC Astacoides; Astacidae; Pacifastacus.
OX NCBI_TaxID=6720;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99162541; Pubmed=10051578;
RA Hall M., Wang R., van Amerongen R., Sottrup-Jensen L., Soderhall K.;
RT "The crayfish plasma clotting protein: a vitellogenin-related protein
RT responsible for clot formation in crustacean blood.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:1965-1970(1999).
DR EMBL; AF102268; AAD16454.1; -.
DR PIR; S30348; S30348.
DR GO; GO:0005319; F:lipid transporter activity; IEA.
DR GO; GO:0006869; F:lipid transport; IEA.
DR InterPro; IPR001747; Lipid_transpt_N.
DR InterPro; IPR01030; LV_superhelical.
DR InterPro; IPR01846; VWF_D.
DR Pfam; PF01347; Vitellogenin_N; 1.
DR Pfam; PF00094; WVD; 1.
DR SMART; SMO0638; LPD_N; 1.
DR SMART; SMO0216; WVD; 1.
DR Signal.
KW Signal.
FT SIGNAL 1
FT CHAIN 16 1721 Potential.
SQ SEQUENCE 1721 AA; 19465 MW; 5BD8CC9352C50CE7 CRC64;

Query Match
Best Local Similarity 76.9%; Score 40; DB 2; Length 1721;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIEYGPLV 10
DB 84 LPIEYGPLV 92

RESULT 12
Q7NNI4 PRELIMINARY; PRT; 3277 AA.
AC Q7NNI4;

```

DT 01-MAR-2004 (TREMBlrel. 26, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE G110427 protein.  
 GN OrderedLocustNames=g110427;  
 OS Gloebacter violaceus.  
 OC Bacteria; Cyanobacteria; Chroococcales; Gloebacter.  
 NC NCB1\_TaxID=33072;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCC 7421;  
 RX MEDLINE=22977040; PubMed=14621292;  
 RA Nakamura Y., Kaneko T., Sato S., Miumuro M., Miyashita H., Tsuchiya T.,  
 RA Koshimizu S., Matsumoto M., Matsuno A., Nakazaki N., Shimo S.,  
 RA Takemuchi C., Yamada M., Tabara S.;  
 RT "Complete genome structure of Gloebacter violaceus PCC 7421. a  
 RT cyanobacterium that lacks thylakoids."  
 RL DNA Res. 10:137-145(2003).  
 DR EMBL; AP06569; BAC88368.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 3277 AA; 337749 MW; B4F90F2EDBBA5CD8 CRC64;

Query Match 76.9%; Score 40; DB 2; Length 3277;  
 Best Local Similarity 75.0%; Pred. No. 4e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 PIRYGLV 10  
 |||||  
 DB 2891 PIRYGLV 2898

RESULT 13  
 Q8RN67 PRELIMINARY; PRT; 144 AA.  
 AC Q8RN67;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=CE2280;  
 OS Corynebacterium efficiens.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
 NC NCB1\_TaxID=152794;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YS-314;  
 RX MEDLINE=22723752; PubMed=12840036;  
 RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,  
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,  
 RA Gotohori T.;  
 RT "Comparative complete genome sequence analysis of the amino acid  
 RT replacements responsible for the thermostability of Corynebacterium  
 RT efficiens."  
 RL Genome Res. 13:1572-1579(2003).  
 DR EMBL; AP065221; BAC19090.1; -.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 144 AA; 16054 MW; 34162E298CAFA4182 CRC64;

Query Match 75.0%; Score 39; DB 2; Length 144;  
 Best Local Similarity 85.7%; Pred. No. 24;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PIRYGLV 9  
 |||||  
 DB 10 PIRYGLV 16

RESULT 14  
 Q8SD08 PRELIMINARY; PRT; 161 AA.  
 ID Q8SD08  
 AC Q8SD08;

DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE PHIKZ154.  
 GN Name=CRF154;  
 OS Pseudomonas phage phikz.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.  
 NC NCB1\_TaxID=169683;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21914557; PubMed=11916376;  
 RA Mesyanzhinov V.V., Robben J., Grymoprez B., Kostyuchenko V.A.,  
 RA Boukaltcheva M.V., Syktilinda N.N., Knylov V.N., Volckart G.;  
 RT "The genome of bacteriophage phikz of Pseudomonas aeruginosa."  
 RL J. Mol. Biol. 317:1-19(2002).  
 DR EMBL; AF39011; ALB83055.1; -.  
 SQ SEQUENCE 161 AA; 18806 MW; 55520C3654704C89 CRC64;

Query Match 73.1%; Score 38; DB 2; Length 161;  
 Best Local Similarity 77.8%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPIRYGLV 10  
 |||||  
 DB 94 LPIRYGLV 102

RESULT 15  
 Q9KRD9 PRELIMINARY; PRT; 163 AA.  
 AC Q9KRD9;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein VC11703.  
 GN OrderedLocustNames=VC11703;  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 NC NCB1\_TaxID=566;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;  
 RA Heidelberg J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
 RA Dodson R.U., Hail D.H., Hickey E.K., Peterson J.D., Nayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,  
 RA Ermolaeva M.D., Vamathevan J.U., Bass S., Qin H., Dragoi I.,  
 RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,  
 RA Newman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,  
 RA Mekalanos J.J., Venter J.C., Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae."  
 RL Nature 406:477-483(2000).  
 DR EMBL; AE004246; AAF94853.1; -.  
 DR F1R; D82167; D82167.  
 DR TIGR; VC11703; -.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 163 AA; 17925 MW; FD33C0A29294B72 CRC64;

Query Match 73.1%; Score 38; DB 2; Length 163;  
 Best Local Similarity 75.0%; Pred. No. 43;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRYGLV 9  
 |||||  
 DB 76 LPIRYGLV 83

Search completed: November 30, 2004, 07:17:30  
 Job time : 79.4103 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 77.1795 Seconds  
(without alignments)  
46.480 Million cell updates/sec

Title: US-10-008-377a-6  
Perfect score: 52  
Sequence: 1 ALPFIEXPLV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 2002273 seqs, 35872929 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23sep04:  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003s:\*  
7: geneseqp2003ps:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	10	AAO22988	AAO22988 Human p45
2	52	100.0	525	AA78521	AA78521 NPW/ALK f
3	52	100.0	525	ABU04368	ABU04368 Human exp
4	52	100.0	572	AA78520	AA78520 Partial A
5	52	100.0	675	ABG20950	ABG20950 Novel hum
6	52	100.0	680	AA873858	AA873858 Human NPM
7	52	100.0	680	ABG95067	ABG95067 Human tra
8	52	100.0	680	ABU04347	ABU04347 Human exp
9	52	100.0	680	ABU04362	ABU04362 Human exp
10	52	100.0	680	ABU04358	ABU04358 Human exp
11	52	100.0	680	ABG74472	ABG74472 NPW/ALK f
12	52	100.0	1620	AA873857	AA873857 Human ALK
13	52	100.0	1620	AAO18498	AAO18498 Human ins
14	52	100.0	1620	ABG74471	ABG74471 Human ALK
15	52	100.0	207	ADFO6379	ADFO6379 Bacterial
16	38	73.1	409	ABO69258	ABO69258 Pseudomon
17	37	71.2	58	AA899175	AA899175 Genetical
18	37	71.2	60	AD771371	AD771371 Kalikre
19	37	71.2	206	ABU04362	ABU04362 LDI recep
20	37	71.2	1083	ABU05679	ABU05679 M. tuberc
21	36	69.2	62	AAU61865	AAU61865 Proionib
22	36	69.2	62	ABMS8384	ABMS8384 Proionib
23	36	69.2	95	ABP05147	ABP05147 Human ORF
24	36	69.2	203	ABRS3343	ABRS3343 Protein s
25	36	69.2	203	ADK63424	ADK63424 Disease t

25	36	69.2	243	7	ABO80522	ABO80522 Pseudomon
27	36	69.2	326	8	ADP29000	ADP29000 C. albica
28	36	69.2	492	4	ABG28689	ABG28689 Novel hum
29	36	69.2	653	2	AA873425	AA873425 IDUA, 3/2
30	36	69.2	728	6	ABU18155	ABU18155 Protein e
31	35	67.3	51	2	AA892258	AA892258 TPPI-2 fi
32	35	67.3	53	2	AA874976	AA874976 Generic f
33	35	67.3	57	3	AA814174	AA814174 Tissue fa
34	35	67.3	58	2	AA878547	AA878547 Human TFP
35	35	67.3	58	2	AA878548	AA878548 Human TFP
36	35	67.3	58	2	AA878549	AA878549 Human TFP
37	35	67.3	58	2	AA878550	AA878550 Human TFP
38	35	67.3	58	2	AA878551	AA878551 Human TFP
39	35	67.3	58	2	AA878552	AA878552 Human TFP
40	35	67.3	58	2	AA878553	AA878553 Human TFP
41	35	67.3	58	2	AA878554	AA878554 Human TFP
42	35	67.3	58	2	AA878555	AA878555 Human TFP
43	35	67.3	58	2	AA878556	AA878556 Human TFP
44	35	67.3	58	2	AA878557	AA878557 Human TFP
45	35	67.3	58	2	AA878558	AA878558 Human TFP

## ALIGNMENTS

RESULT 1	AAO22988	standard; peptide; 10 AA.
ID	AAO22988	
AC	AAO22988	
XX	AAO22988	
DT	17-SEP-2003	(first entry)
XX	17-SEP-2003	
DE	Human p45e-65 ALK-derived HLA-A*0201 restricted CTL epitope peptide.	
XX	HLA-A*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;	
KM	cytotoxic T-lymphocyte receptor tyrosine; NPW/ALK fusion protein; human;	
KM	oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;	
KM	t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;	
KM	p45e-65.	
OS	Homo sapiens.	
XX	MO2003042243-A2.	
PN	MO2003042243-A2.	
XX	22-MAY-2003.	
PD	22-MAY-2003.	
XX	14-NOV-2002; 2002WO-EP012764.	
PF	14-NOV-2002; 2002WO-EP012764.	
XX	15-NOV-2001; 2001US-00008377.	
PR	15-NOV-2001; 2001US-00008377.	
XX	(NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.	
PA	(NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.	
XX	(STUD-) ST JUDE CHILDREN'S RES HOSPITAL.	
XX	Gambacorti-Passerini C, Passoni L;	
PI	Gambacorti-Passerini C, Passoni L;	
XX	WPI, 2003-441791/41.	
DR	WPI, 2003-441791/41.	
XX	New HLA-A*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,	
PT	useful for preparing a composition for treating ALK-positive lymphoma,	
PT	neuroblastoma or ALK-expressing neoplasia.	
XX	Claim 1, Page 7, 33pp; English.	
PS	Claim 1, Page 7, 33pp; English.	
XX	The invention relates to a novel HLA-A*0201-binding anaplastic lymphoma	
CC	kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a	
CC	cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase	
CC	which forms part of the NPW/ALK fusion protein (oncogenic lymphoma kinase	
CC	fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell	
CC	lines. More than 50% of ALCL cases possess a t(2;5) chromosomal	
CC	translocation that leads to the expression of the NPW/ALK fusion protein	
CC	which forms a potent oncogene when constitutively activated. Translocated	
CC	ALK is a widely expressed tumour-associated antigen characteristic of ALK	

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p456-65  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 XX

SQ Sequence 10 AA;

Query Match 100.0%; Score 52; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0064;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALPIEYGPLV 10  
 |||||  
 Db 1 ALPIEYGPLV 10

RESULT 2  
 AAR78521  
 ID AAR78521 standard; protein; 525 AA.  
 XX  
 AC AAR78521;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 15-NOV-1995 (first entry)  
 XX  
 DE NPM/ALK fusion protein.  
 XX  
 KW ALK protein; protein-tyrosine-kinase; fusion protein; NPM protein;  
 KW nucleolar phosphoprotein; anaplastic large cell lymphoma; t(2;  
 XX 5) translocation; diagnosis; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO9515331-A1.  
 XX  
 PD 08-JUN-1995.  
 XX  
 PF 05-DEC-1994; 94WO-US013947.  
 XX  
 PR 03-DEC-1993; 93US-00160861.  
 XX  
 PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Morris SW, Look AT;  
 XX  
 DR WPI; 1995-215226/28.  
 DR N-PEDB; AAQ95542.  
 XX  
 PT Methode for detecting human t(2;5) lymphoma - for detection and diagnosis  
 PT of anaplastic large cell lymphoma(e).  
 XX  
 PS Claim 16; Page 42-44; 70pp; English.  
 XX  
 CC The translocation event that occurs in human t(2;5) lymphoma brings  
 CC sequences from the nucleolar phosphoprotein (NPM) gene on chromosome 5q35  
 CC to those from a protein-tyrosine-kinase (ALK) gene on chromosome-2q23.  
 CC Nucleic acids encoding a partial ALK protein (AAR78520) and the ALK/NPM  
 CC fusion protein (AAR78521) were isolated (AAQ95541, AAQ95542).  
 CC Identification of the NPM/ALK fusion allows lymphoma diagnosis, or  
 CC therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR  
 CC -2003 to correct PN field.)  
 XX  
 SQ Sequence 525 AA;

Query Match 100.0%; Score 52; DB 2; Length 525;  
 Best Local Similarity 100.0%; Pred. No. 0.49;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALPIEYGPLV 10  
 |||||  
 Db 456 ALPIEYGPLV 465

RESULT 3  
 ABU04368  
 ID ABU04368 standard; protein; 525 AA.  
 XX  
 AC ABU04368;  
 XX  
 DT 29-JAN-2003 (first entry)  
 XX  
 DE Human expressed protein tag (EPT) #1034.  
 XX  
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200278524-A2.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 28-MAR-2002; 2002WO-US009671.  
 XX  
 PR 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 PA (ZYCO-) ZYCOs INC.  
 XX  
 PI Chicz RM, Tomlinson AJ, Urban RG;  
 XX  
 DR WPI; 2003-040607/03.  
 XX  
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX  
 XX Example 2; SEQ ID NO 1034; 134pp; English.  
 XX  
 PS The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 525 AA;

Query Match 100.0%; Score 52; DB 6; Length 525;  
 Best Local Similarity 100.0%; Pred. No. 0.49;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALPIEXGPLV 10  
 |||||  
 DB 456 ALPIEXGPLV 465

## RESULT 4

AAR78520 standard; protein; 572 AA.

AC AAR78520;

DT 25-MAR-2003 (revised)  
 DT 15-NOV-1995 (first entry)

DE Partial ALK protein.

KM ALK protein; protein-tyrosine-kinase; fusion protein; NPM protein;  
 KM nuclear phosphoprotein; anaplastic large cell lymphoma; t(2;  
 KM 5) translocation; diagnosis; gene therapy.

OS Homo sapiens.

PN WO9515331-A1.

PD 08-JUN-1995.

PF 05-DEC-1994; 94WO-US013947.

PR 03-DEC-1993; 93US-00160861.

PA (SUUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Morris SW, Look AT;

XX WPI; 1995-215226/28.

DR N-PSDB; AAQ95541.

PT Methods for detecting human t(2;5) lymphoma - for detection and diagnosis

XX of anaplastic large cell lymphoma(s).

PS Claim 21; Page 40-42; 70pp; English.

XX The translocation event that occurs in human t(2;5) lymphoma brings

CC sequences from the nucleolar phosphoprotein (NPM) gene on chromosome 5q35

CC to those from a protein-tyrosine-kinase (ALK) gene on chromosome 2q32.

CC Nucleic acids encoding a partial ALK protein (AAR78520) and the ALK/NPM

CC fusion protein (AAR78521) were isolated (AAQ95541, AAQ95542).

CC Identification of the NPM/ALK fusion allows lymphoma diagnosis, or

CC therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR

CC -2003 to correct PN field.)

XX

XX Sequence 572 AA;

XX

XX

XX

XX

XX

XX

XX

KM food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US008631.

PF 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.

DR N-PSDB; AAS85137.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

OS Homo sapiens.  
 XX US6174674-B1.  
 XX 16-JAN-2001.  
 PD 19-JUN-1998; 98US-00100089.  
 XX 03-DEC-1993; 93US-00160861.  
 PR 12-OCT-1995; 95US-00542363.  
 XX (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 PA Morris SW, Look AT;  
 PI WPI; 2001-243208/25.  
 DR N-PSDB; AAF76868.  
 XX  
 PT Detection of chromosomal rearrangement or translocations present in t(2;  
 PT 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 PT techniques.  
 XX  
 PS Claim 1; Fig 2A; 87pp; English.  
 CC The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 CC  
 XX  
 SQ Sequence 680 AA;

Query Match 100.0%; Score 52; DB 4; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 0.65;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10  
 |||||  
 DB 456 ALPIEYGPLV 465

RESULT 7  
 ID ABG95067 standard; protein; 680 AA.  
 AC ABG95067;

DI 04-DEC-2002 (first entry)

DE Human translocation (2; 5) (p23; q35) protein.

XX Chromosome aberration; oncogenic fusion protein; cancer;  
 XX Proliferative disease; cellular protein isoform; heat shock protein 90;  
 KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;  
 KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;  
 KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;  
 KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;  
 KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;  
 KW rhabdomyosarcoma; synovial sarcoma; viral infection.

OS Homo sapiens.

XX WO200269900-A2.

PD 12-SEP-2002.

PF 01-MAR-2002; 2002WO-US006518.  
 XX 01-MAR-2001; 2001US-0272751P.  
 XX (CONF-) CONFORMA THERAPEUTICS CORP.  
 PA Fritz LC, Burrows FJ;  
 PI WPI; 2002-698716/75.  
 DR N-PSDB; ABS73246.  
 XX  
 PT Treating genetically-defined disease associated with chromosomal  
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative  
 PT diseases, involves administering an inhibitor of heat shock protein 90.  
 XX  
 PS Disclosure; Page 200-202; 389pp; English.  
 CC The invention describes a method of treating genetically-defined disease  
 CC associated with chromosomal aberrations yielding oncogenic fusion  
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous  
 CC cell population, treating proliferative diseases associated with mutant  
 CC protein or cellular protein isoforms (II) dependent on heat shock protein  
 CC (HSP)-90, or selectively treating cells expressing (II) involving  
 CC administering HSP90-inhibitor. The method is useful for treating  
 CC genetically-defined disease with chromosomal aberration yielding  
 CC oncogenic fusion protein, treating cancerous cells containing fusion  
 CC protein in heterogeneous cell population, treating proliferative disease  
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or  
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.  
 CC p53), or selectively treating cells expressing mutant protein or cellular  
 CC protein isoform in a patient heterozygous for (II). The method is useful  
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell  
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML,  
 CC or a disease characterised by a solid tumour such as papillary thyroid  
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and  
 CC synovial sarcoma. The method is also useful for treating viral  
 CC infections. This represents a protein encoded by the DNA sequence of a  
 CC chromosome aberration  
 CC  
 XX  
 SQ Sequence 680 AA;

Query Match 100.0%; Score 52; DB 5; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 0.65;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10  
 |||||  
 DB 456 ALPIEYGPLV 465

RESULT 8  
 ID ABU04347 standard; protein; 680 AA.  
 AC ABU04347;

DI 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1013.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

OS Homo sapiens.

XX WO200278524-A2.

PD 10-OCT-2002.

PF 28-MAR-2002; 2002WO-US009671.

```
XX 28-MAR-2001; 2001US-0279495F.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG,
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1013; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 680 AA;
SQ
XX
XX Query Match 100.0%; Score 52; DB 6; Length 680;
XX Best Local Similarity 100.0%; Pred. No. 0.65;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ALPIEYGPV 10
XX |||||
XX 456 ALPIEYGPV 465
XX
XX Db
XX
XX RESULT 9
XX ABU04362
XX ID ABU04362 standard; protein; 680 AA.
XX
XX AC ABU04362;
XX
XX DT 29-JAN-2003 (first entry)
XX
XX DE Human expressed protein tag (EPT) #1028.
XX
XX KW Translational profiling; expressed protein tag; EPT; Kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.
XX
XX OS Homo sapiens.
XX
XX PN WO200278524-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 28-MAR-2002; 2002WO-US009671.
```

```
XX 28-MAR-2001; 2001US-0279495F.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG,
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1028; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 680 AA;
SQ
XX
XX Query Match 100.0%; Score 52; DB 6; Length 680;
XX Best Local Similarity 100.0%; Pred. No. 0.65;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ALPIEYGPV 10
XX |||||
XX 456 ALPIEYGPV 465
XX
XX Db
XX
XX RESULT 10
XX ABU04358
XX ID ABU04358 standard; protein; 680 AA.
XX
XX AC ABU04358;
XX
XX DT 29-JAN-2003 (first entry)
XX
XX DE Human expressed protein tag (EPT) #1024.
XX
XX KW Translational profiling; expressed protein tag; EPT; Kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.
XX
XX OS Homo sapiens.
XX
XX PN WO200278524-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 28-MAR-2002; 2002WO-US009671.
```

XX 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCO INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.

XX Example 2; SEQ ID NO 1024; 134bp; English.

XX The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 680 AA;

XX Query Match 100.0%; Score 52; DB 6; Length 680;

XX Best Local Similarity 100.0%; Pred. No. 0.65;

XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10

DB 456 ALPIEYGPLV 465

XX RESULT 11

XX AAG74472

XX ID AAG74472 standard; protein; 680 AA.

XX AAG74472;

XX 11-APR-2003 (first entry)

XX NPM/ALK fusion construct SEQ ID 4.

XX ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;

XX t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;

XX nucleolar phosphoprotein; centromeric; telomeric.

XX Synthetic.

XX Key

XX Location/Qualifiers

XX 1..117

XX /note= "NPM-derived protein fragment"

XX 104..115

XX /note= "potential metal binding domain"

XX 118..680

XX /note= "ALK-derived protein fragment"

FT Domain 182..437

FT /label= ALK\_catalytic\_domain

XX US6451997-B1.

XX 17-SEP-2002.

XX 28-SEP-2000; 2000US-00670827.

XX 03-DEC-1993; 93US-00160861.

XX 12-OCT-1995; 95US-00542363.

XX 19-JUN-1998; 98US-00100089.

XX (SUJD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Morris SW, Look AT;

XX WPI; 2003-101739/09.

XX N-PSDB; ABQ77107.

XX Kit for use in method of detecting t(2;5) chromosomal rearrangements or

XX rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic

XX lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK

XX genes.

XX Example 1; Fig 2A; 87bp; English.

XX This invention describes a novel kit for use in a method of detecting t(2

XX /5) chromosomal rearrangements (CR) or CRs involving nucleolar

XX phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method

XX comprises labelling probes that hybridise to a sequence of 2442 or 6226

XX nucleotides or their complement, and to region of human chromosome 5/2

XX having wild-type NPM/ALK genes (the region is centromeric/telomeric to

XX nucleotide). This sequence represents a polypeptide sequence described in

XX the disclosure of the invention

XX Sequence 680 AA;

XX Query Match 100.0%; Score 52; DB 6; Length 680;

XX Best Local Similarity 100.0%; Pred. No. 0.65;

XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10

DB 456 ALPIEYGPLV 465

XX RESULT 12

XX AAB73857

XX ID AAB73857 standard; protein; 1620 AA.

XX AAB73857;

XX 15-MAY-2001 (first entry)

XX Human ALK protein.

XX Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;

XX chromosomal translocation; cancer.

XX Homo sapiens.

XX US6174674-B1.

XX 16-JAN-2001.

XX 19-JUN-1998; 98US-00100089.

XX 03-DEC-1993; 93US-00160861.

XX 12-OCT-1995; 95US-00542363.

XX (SUJD-) ST JUDE CHILDREN'S RES HOSPITAL.



XX Morris SW, Look AT;  
XX MPI, 2001-243208/25.  
DR N-PSDB; AAF76867.  
XX  
XX  
PT Detection of chromosomal rearrangement or translocations present in t(2;  
PT 5) for cancer diagnosis comprises using and identifying anaplastic  
PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
PT techniques.  
XX  
PS Claim 10; Fig 3B; 87pp; English.  
CC  
CC The present sequence is given in a specification relating to a method for  
CC detecting a chromosomal rearrangement involving a breakpoint in the  
CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
CC is based on the identification and sequence determination of a novel  
CC gene, ALK, which is fused to the gene encoding NPM in translocations  
CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
CC involving the ALK gene or NPM gene. It is particularly useful in  
CC determining if particular cells or tissues express ALK or NPM/ALK coding  
CC sequences. The method is also useful in diagnostic assays to determine,  
CC for example, if a mammal has cancer or a genetic predisposition to cancer.  
XX  
SQ Sequence 1620 AA;  
Query Match 100.0%; Score 52; DB 4; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ALPIEXGPLV 10  
Db 1396 ALPIEXGPLV 1405  
RESULT 13  
ID AAO18498  
AAO18498 standard; protein; 1620 AA.  
XX  
XX AAO18498;  
XX  
DT 11-OCT-2002 (first entry)  
DE  
DE Human insulin receptor signaling modifier SEQ ID NO: 18.  
KW Human; insulin receptor signaling; insulin receptor signaling modifier;  
KW IHW; diabetes; metabolic syndrome; antidiabetic.  
XX  
XX Homo sapiens.  
OS  
PN WO200255664-A2.  
XX  
PD 18-JUL-2002.  
XX  
PF 11-JAN-2002; 2002WO-US001048.  
XX  
PR 12-JAN-2001; 2001US-0261226C.  
PR 12-JAN-2001; 2001US-0261303P.  
PR 12-JAN-2001; 2001US-0261304P.  
PR 12-JAN-2001; 2001US-0261335P.  
PR 12-JAN-2001; 2001US-0261336P.  
PR 12-JAN-2001; 2001US-0261361P.  
PR 12-JAN-2001; 2001US-0261456P.  
PR 12-JAN-2001; 2001US-0261457P.  
PR 12-JAN-2001; 2001US-0261458P.  
PR 12-JAN-2001; 2001US-0261459P.  
PR 12-JAN-2001; 2001US-0261461P.  
PR 12-JAN-2001; 2001US-0261518P.  
PR 12-JAN-2001; 2001US-0261531P.  
PR 12-JAN-2001; 2001US-0261532P.  
PR 12-JAN-2001; 2001US-0261589P.  
PR 12-JAN-2001; 2001US-0261590P.  
PR 12-JAN-2001; 2001US-0261694P.  
DR

Query Match	100.0%	Score 52	DB 5	Length 1620
Best Local Similarity	100.0%	Pred. No. 1.7		
Matches 10	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1 ALPIEYGPLV 10			
DB	1396 ALPIEYGPLV 1405			
RESULT 14				
ID	ABG74471			
XX	ABG74471 standard; protein; 1620 AA.			
AC	ABG74471:			
DT	11-APR-2003 (first entry)			
XX				
DE	Human ALK protein SEQ ID 2.			
XX				
KW	ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;			
KW	t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;			
KM	nucleolar phosphoprotein; centromeric; telomeric.			
XX				
OS	Homo sapiens.			
XX				
Key	Location/Qualifiers			
FT	Peptide 1..26			
FT	/label= signal_peptide			
FT	Protein 27..1620			
FT	/label= mature_alk			
FT	Domain 1031..1058			
FT	/label= transmembrane_domain			
FT	Misc-difference 1058..1059			
FT	/note= "NPM-ALK fusion junction"			
FT	Domain 1123..1377			
FT	/label= tyrosine_kinase_catalytic_domain			
XX				
XX	US6451997-B1.			
XX				
XX	17-SEP-2002.			

XX 28-SEP-2000; 2000US-00670827.  
 XX  
 PR 03-DEC-1993; 93US-00160861.  
 PR 12-OCT-1995; 95US-00542363.  
 PR 19-JUN-1998; 98US-00100089.  
 XX  
 XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 PA  
 PI Morris SW, Look AT;  
 XX  
 DR WPI; 2003-101739/09.  
 DR N-PSDB; ABQ77106.  
 PT  
 PT Kit for use in method of detecting t(2;5) chromosomal rearrangements or  
 PT rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic  
 PT lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK  
 PT genes.  
 XX  
 XX  
 PS Example 1; Fig 3B; 87pp; English.  
 XX  
 CC This invention describes a novel kit for use in a method of detecting t(2  
 CC :5) chromosomal rearrangements (CR) or CRs involving nucleolar  
 CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method  
 CC comprises labelled probes that hybridise to a sequence of 2442 or 6226  
 CC nucleotides or their complement, and to region of human chromosome 5/2  
 CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to  
 CC nucleotide). This sequence represents a polypeptide sequence described in  
 CC the disclosure of the invention  
 XX  
 SQ Sequence 1620 AA;

Query Match 100.0%; Score 52; DB 6; Length 1620;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALPIRYGPLYV 10  
 DB 1396 ALPIRYGPLYV 1405

RESULT 15  
 ADF06379  
 ID ADF06379 standard; protein; 207 AA.  
 XX  
 AC ADF06379;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Bacterial polypeptide #2492.  
 XX  
 XX Proteus mirabilis infection; bacterial infection; antibacterial;  
 KW immunostimulant.  
 XX  
 OS Proteus mirabilis.  
 XX  
 PN US6605709-B1.  
 XX  
 PD 12-AUG-2003.  
 XX  
 PF 05-APR-2000; 2000US-00543681.  
 XX  
 PR 09-APR-1999; 99US-0128706P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Breton GL;  
 XX  
 DR WPI; 2003-895291/82.  
 DR N-PSDB; ADF02207.  
 XX  
 PT New Proteus mirabilis polypeptides and polynucleotides, useful as  
 PT reagents for diagnosis of bacterial disease, as components of

PT antibacterial vaccines, as targets for antibacterial drugs, or as  
 PT biocontrol agents for plants.  
 PS  
 XX Disclosure; SEQ ID NO 6664; 870pp; English.  
 CC  
 CC The invention relates to new Proteus mirabilis polypeptides and  
 CC polynucleotides. The invention also relates to antibodies against the  
 CC polynucleotides, methods for producing the polypeptides, a method of  
 CC generating vaccines for immunising an individual against P. mirabilis, a  
 CC method for evaluating a compound for the ability to bind a P. mirabilis  
 CC polypeptide and a method for screening test compounds for anti-bacterial  
 CC activity. The polypeptides and polynucleotides are useful as molecular  
 CC targets for diagnosing, preventing and treating pathological conditions  
 CC resulting from bacterial infection, as reagents for diagnosis of  
 CC bacterial diseases, as components of antibacterial vaccines, as targets  
 CC for antibacterial drugs or as bio-control agents for plants. This  
 CC sequence represents a Proteus mirabilis polypeptide of the invention.  
 XX

Query Match 73.1%; Score 38; DB 7; Length 207;  
 Best Local Similarity 60.0%; Pred. No. 73;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ALPIRYGPLYV 10  
 DB 110 ALDMEYGPII 119

Search completed: November 30, 2004, 07:39:57  
 Job time : 79.4295 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 30, 2004, 07:19:22 ; Search time 376.923 Seconds  
(without alignments)  
9,410 Million cell updates/sec

Title: US-10-008-377a-6

Perfect score: 52

Sequence: 1 ALPIEXGPLV 10

## Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US06\_PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	14	US-10-008-377-6
2	52	100.0	370	16	US-10-664-421-70
3	52	100.0	680	9	US-09-827-949-4
4	52	100.0	1620	9	US-09-827-949-2
5	39	75.0	121	15	US-10-424-599-264129
6	39	75.0	685	15	US-10-425-114-45829
7	38	73.1	401	15	US-10-389-647-586
8	37	71.2	58	14	US-10-038-722-34
9	37	71.2	60	14	US-10-456-966A-51
10	37	71.2	206	14	US-10-211-962-55
11	37	71.2	677	14	US-10-425-114-45922
12	37	71.2	1083	14	US-10-080-170-330
13	37	71.2	1083	16	US-10-080-170-330

14	37	71.2	1083	17	US-10-468-356-330	Sequence 330, App
15	36	69.2	287	15	US-10-424-599-213533	Sequence 213533
16	36	69.2	446	17	US-10-425-115-267156	Sequence 267156
17	36	69.2	695	16	US-10-437-966-180466	Sequence 180466
18	36	69.2	728	15	US-10-282-122A-46079	Sequence 46079, A
19	35	67.3	51	9	US-09-827-948-7	Sequence 7, Appl1
20	35	67.3	51	9	US-09-741-106-2	Sequence 7, Appl1
21	35	67.3	51	14	US-10-176-071-7	Sequence 7, Appl1
22	35	67.3	55	9	US-09-904-621-15	Sequence 15, Appl1
23	35	67.3	57	10	US-09-974-026-21	Sequence 21, Appl1
24	35	67.3	58	14	US-10-167-351-70	Sequence 70, Appl1
25	35	67.3	58	14	US-10-038-722-32	Sequence 32, Appl1
26	35	67.3	58	15	US-10-456-966A-37	Sequence 37, Appl1
27	35	67.3	58	16	US-10-361-997-12	Sequence 12, Appl1
28	35	67.3	61	16	US-10-767-701-51101	Sequence 51101, A
29	35	67.3	78	16	US-10-799-326-71	Sequence 71, Appl1
30	35	67.3	88	10	US-09-978-418-26	Sequence 26, Appl1
31	35	67.3	123	14	US-10-356-088-1	Sequence 1, Appl1
32	35	67.3	123	14	US-10-356-088-23	Sequence 23, Appl1
33	35	67.3	123	14	US-10-353-454-1	Sequence 1, Appl1
34	35	67.3	123	14	US-10-353-454-23	Sequence 23, Appl1
35	35	67.3	123	16	US-10-799-326-1	Sequence 1, Appl1
36	35	67.3	123	16	US-10-799-326-23	Sequence 23, Appl1
37	35	67.3	124	15	US-10-425-156-10	Sequence 10, Appl1
38	35	67.3	124	15	US-09-766-778-2	Sequence 2, Appl1
39	35	67.3	213	9	US-10-086-176A-6	Sequence 6, Appl1
40	35	67.3	235	9	US-09-904-621-2	Sequence 2, Appl1
41	35	67.3	235	9	US-09-736-457-332	Sequence 332, App
42	35	67.3	235	9	US-09-902-941-332	Sequence 332, App
43	35	67.3	235	9	US-09-849-300-332	Sequence 332, App
44	35	67.3	235	10	US-09-476-360-332	Sequence 332, App
45	35	67.3	235	14	US-10-017-754-332	Sequence 332, App

## ALIGNMENTS

RESULT 1

US-10-008-377-6

Sequence 6, Appl1 Application US/10008377

Publication No. US2003015101A1

GENERAL INFORMATION:

APPLICANT: Gambacorti-Passerini, Carlo

APPLICANT: Passoni, Lorena

TITLE OF INVENTION: Immunogenic ALK Peptides

FILE REFERENCE: 045922/241203

CURRENT APPLICATION NUMBER: US/10/008,377

CURRENT FILING DATE: 2001-11-15

NUMBER OF SEQ ID NOS: 7

SOFTWARE: FastSeq For Windows Version 4.0

SEQ ID NO: 6

LENGTH: 10

TYPE: PRT

ORGANISM: Homo sapiens

US-10-008-377-6

Query Match

Best Local Similarity 100.0%; Score 52; DB 14; Length 10;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEXGPLV 10

DB 1 ALPIEXGPLV 10

RESULT 2

US-10-664-421-70

Sequence 70, Appl1 Application US/10064421

Publication No. US20040142854A1

GENERAL INFORMATION:

APPLICANT: BREWER, RYAN

APPLICANT: IBERAHIM, PRAHBA

APPLICANT: KUMAR, ABHINAV

```

; APPLICANT: MANDIYAN, VALISAN
; APPLICANT: MILBURN, MICHAEL V.
; TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE
; FILE REFERENCE: 039363/0703
; CURRENT APPLICATION NUMBER: US/10/664,421
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/412,341
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/411,398
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 70
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-664-421-70

```

```

Query Match      100.0%; Score 52; DB 16; Length 370;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ALPIEYGPLV 10
      |||||
Db      331 ALPIEYGPLV 340

```

```

RESULT 3
US-09-827-949-4
; Sequence 4, Application US/09827949
; Patent No. US20010021505A1
; GENERAL INFORMATION:
; APPLICANT: MORRIS, STEPHAN W.
; APPLICANT: LOCK, A. THOMAS
; TITLE OF INVENTION: ALK PROTEIN TYROSINE KINASE/RECEPTOR AND LIGANDS THEREOF
; FILE REFERENCE: 0656.0400004
; CURRENT APPLICATION NUMBER: US/09/827,949
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 09/670,827
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 09/100,089
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: US 08/542,363
; PRIOR FILING DATE: 1995-10-12
; PRIOR APPLICATION NUMBER: US 08/160,861
; PRIOR FILING DATE: 1993-12-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 680
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-827-949-4

```

```

Query Match      100.0%; Score 52; DB 9; Length 680;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ALPIEYGPLV 10
      |||||
Db      456 ALPIEYGPLV 465

```

```

RESULT 4
US-09-827-949-2
; Sequence 2, Application US/09827949
; Patent No. US20010021505A1
; GENERAL INFORMATION:
; APPLICANT: MORRIS, STEPHAN W.
; APPLICANT: LOCK, A. THOMAS
; TITLE OF INVENTION: ALK PROTEIN TYROSINE KINASE/RECEPTOR AND LIGANDS THEREOF
; FILE REFERENCE: 0656.0400004
; CURRENT APPLICATION NUMBER: US/09/827,949

```

```

; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 09/670,827
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 09/100,089
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: US 08/542,363
; PRIOR FILING DATE: 1995-10-12
; PRIOR APPLICATION NUMBER: US 08/160,861
; PRIOR FILING DATE: 1993-12-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-827-949-2

```

```

Query Match      100.0%; Score 52; DB 9; Length 1620;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ALPIEYGPLV 10
      |||||
Db      1396 ALPIEYGPLV 1405

```

```

RESULT 5
US-10-424-599-264129
; Sequence 264129, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: LA ROSA, THOMAS J
; APPLICANT: KOWALIC, DAVID K
; APPLICANT: ZHOU, YIHUA
; APPLICANT: CAO, YONGWEI
; TITLE OF INVENTION: SOY NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53423)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 265684
; SEQ ID NO 264129
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: CLONE ID: PAT_MRT3847_805290.1.pep
US-10-424-599-264129

```

```

Query Match      75.0%; Score 39; DB 15; Length 121;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 LPIEYGPLV 10
      |||||
Db      94 LPIEYGPLV 102

```

```

RESULT 6
US-10-425-114-45829
; Sequence 45829, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: LIU, JINGDONG
; APPLICANT: ZHOU, YIHUA
; APPLICANT: KOWALIC, DAVID K.
; APPLICANT: SCREEN, STEVEN E
; APPLICANT: TABASKA, JACK E
; APPLICANT: CAO, YONGWEI
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53413)B
; CURRENT APPLICATION NUMBER: US/10/425,114

```

;; CURRENT FILING DATE: 2003-04-28  
;; NUMBER OF SEQ ID NOS: 73128  
;; SEQ ID NO 45829  
;; LENGTH: 685  
;; TYPE: PRT  
;; ORGANISM: Glycine max  
;; FEATURE:  
;; OTHER INFORMATION: Clone ID: 700910011\_Flt.pap  
US-10-425-114-45829

Query Match 75.0%; Score 39; DB 15; Length 685;  
Best Local Similarity 60.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALPIRYGPLY 10  
DB 147 SLPLKXGPFV 156

RESULT 7  
US-10-389-647-586  
;; Sequence 586, Application US/10389647  
;; Publication No. US20040033549A1  
;; GENERAL INFORMATION:  
;; APPLICANT: GREENBERG, E. Peter  
;; APPLICANT: SCHUSTER, Martin  
;; APPLICANT: LOSTROH, Carol  
;; TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA  
;; FILE REFERENCE: U2-038CP  
;; CURRENT APPLICATION NUMBER: US/10/389,647  
;; CURRENT FILING DATE: 2003-03-14  
;; PRIOR APPLICATION NUMBER: 09/653730  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/153022  
;; PRIOR FILING DATE: 1999-09-03  
;; NUMBER OF SEQ ID NOS: 710  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 586  
;; LENGTH: 401  
;; TYPE: PRT  
;; ORGANISM: Pseudomonas aeruginosa  
US-10-389-647-586

Query Match 73.1%; Score 38; DB 15; Length 401;  
Best Local Similarity 88.9%; Pred. No. 1.6e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALPIRYGPLY 9  
DB 278 ALPIRYGPLY 286

RESULT 8  
US-10-038-722-34  
;; Sequence 34, Application US/10038722  
;; Publication No. US20030175919A1  
;; GENERAL INFORMATION:  
;; APPLICANT: LEY, Arthur C.  
;; APPLICANT: GUTERMAN, Sonia K.  
;; APPLICANT: MARKLAND, William  
;; APPLICANT: KENT, Rachel B.  
;; APPLICANT: ROBERTS, Bruce L.  
;; APPLICANT: LADNER, Robert C.  
;; TITLE OF INVENTION: ITI-D1 KUNITZ DOMAIN MUTANTS AS NHE INHIBITORS  
;; FILE REFERENCE: LEY-1B  
;; CURRENT APPLICATION NUMBER: US/10/038,722  
;; CURRENT FILING DATE: 2002-01-08  
;; PRIOR APPLICATION NUMBER: US 08/649,406  
;; PRIOR FILING DATE: 1999-07-21  
;; PRIOR APPLICATION NUMBER: PCT/US95/16349  
;; PRIOR FILING DATE: 1995-12-15  
;; PRIOR APPLICATION NUMBER: US 08/358,160  
;; PRIOR FILING DATE: 1994-12-16

;; NUMBER OF SEQ ID NOS: 129  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 34  
;; LENGTH: 58  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: DPL 2.2  
US-10-038-722-34

Query Match 71.2%; Score 37; DB 14; Length 58;  
Best Local Similarity 55.6%; Pred. No. 32;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIRYGPLY 10  
DB 7 LPLDYGPFI 15

RESULT 9  
US-10-456-986A-51  
;; Sequence 51, Application US/10456986A  
;; Publication No. US2004003893A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ladner, Robert C.  
;; APPLICANT: Ley, Arthur C.  
;; APPLICANT: HIRANI, Shifish  
;; APPLICANT: Williams, Anthony  
;; TITLE OF INVENTION: Prevention and Reduction of Blood Loss  
;; FILE REFERENCE: 3421.1001-002  
;; CURRENT APPLICATION NUMBER: US/10/456,986A  
;; CURRENT FILING DATE: 2003-06-06  
;; PRIOR APPLICATION NUMBER: 60/387,239  
;; PRIOR FILING DATE: 2002-06-07  
;; PRIOR APPLICATION NUMBER: 60/407,003  
;; PRIOR FILING DATE: 2002-08-28  
;; NUMBER OF SEQ ID NOS: 54  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 51  
;; LENGTH: 60  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: DPL24 KR Sequence  
US-10-456-986A-51

Query Match 71.2%; Score 37; DB 15; Length 60;  
Best Local Similarity 55.6%; Pred. No. 32;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIRYGPLY 10  
DB 9 LPLDYGPFI 17

RESULT 10  
US-10-211-962-55  
;; Sequence 55, Application US/10211962  
;; Publication No. US20030082640A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Herz, Joachim  
;; APPLICANT: Gotthardt, Michael  
;; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
;; FILE REFERENCE: U2000708  
;; CURRENT APPLICATION NUMBER: US/10/211,962  
;; CURRENT FILING DATE: 2002-08-01  
;; PRIOR APPLICATION NUMBER: US/08/562,737  
;; PRIOR FILING DATE: 2000-05-01  
;; NUMBER OF SEQ ID NOS: 132  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 55  
;; LENGTH: 206  
;; TYPE: PRT

ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 OTHER INFORMATION: Sequence  
 US-10-211-962-55

Query Match 71.2%; Score 37; DB 14; Length 206;  
 Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEYGPLV 10  
 DB 161 LPVQMGPIV 169

RESULT 11  
 US-10-425-114-45922  
 Sequence 45922, Application US/10425114  
 Publication No. US20040034888A1  
 GENERAL INFORMATION:  
 APPLICANT: Liu, Jingdong  
 APPLICANT: Zhou, Yinhua  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Screen, Steven E  
 APPLICANT: Tabaska, Jack E  
 APPLICANT: Gao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(3313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 45922  
 LENGTH: 677  
 TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 OTHER INFORMATION: Clone ID: 701011172\_FLI.pep  
 US-10-425-114-45922

Query Match 71.2%; Score 37; DB 15; Length 677;  
 Best Local Similarity 55.6%; Pred. No. 4.1e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIEYGPLV 10  
 DB 142 LPLKXGPI 150

RESULT 12  
 US-10-080-170-330  
 Sequence 330, Application US/10080170  
 Publication No. US20030129601A1  
 GENERAL INFORMATION:  
 APPLICANT: COLE, S.T.  
 TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR  
 TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR  
 FILE REFERENCE: 03495,0218  
 CURRENT APPLICATION NUMBER: US/10/080,170  
 CURRENT FILING DATE: 2002-06-10  
 PRIOR APPLICATION NUMBER: 60/270,123  
 PRIOR FILING DATE: 2001-02-22  
 NUMBER OF SEQ ID NOS: 652  
 SOFTWARE: Patent Ver. 2.1  
 SEQ ID NO 330  
 LENGTH: 1083  
 TYPE: PRT  
 ORGANISM: Mycobacterium leprae  
 US-10-080-170-330

Query Match 71.2%; Score 37; DB 14; Length 1083;  
 Best Local Similarity 77.8%; Pred. No. 6.7e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 ALPIEYGPL 9  
 DB 753 ALPSNYGPL 761

RESULT 13  
 US-10-080-170-330  
 Sequence 330, Application US/10080170  
 Publication No. US20040121322A9  
 GENERAL INFORMATION:  
 APPLICANT: COLE, S.T.  
 TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR  
 TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR  
 FILE REFERENCE: 03495,0218  
 CURRENT APPLICATION NUMBER: US/10/080,170  
 CURRENT FILING DATE: 2002-06-10  
 PRIOR APPLICATION NUMBER: 60/270,123  
 PRIOR FILING DATE: 2001-02-22  
 NUMBER OF SEQ ID NOS: 652  
 SOFTWARE: Patent Ver. 2.1  
 SEQ ID NO 330  
 LENGTH: 1083  
 TYPE: PRT  
 ORGANISM: Mycobacterium leprae  
 US-10-080-170-330

Query Match 71.2%; Score 37; DB 16; Length 1083;  
 Best Local Similarity 77.8%; Pred. No. 6.7e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ALPIEYGPL 9  
 DB 753 ALPSNYGPL 761

RESULT 14  
 US-10-468-356-330  
 Sequence 330, Application US/10468356  
 Publication No. US20040197896A1  
 GENERAL INFORMATION:  
 APPLICANT: COLE, STEWART  
 TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR  
 TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR  
 FILE REFERENCE: 03394,0019  
 CURRENT APPLICATION NUMBER: US/10/468,356  
 CURRENT FILING DATE: 2003-08-19  
 PRIOR APPLICATION NUMBER: 10/080,170  
 PRIOR FILING DATE: 2002-02-22  
 PRIOR APPLICATION NUMBER: 60/270,123  
 PRIOR FILING DATE: 2001-02-22  
 NUMBER OF SEQ ID NOS: 655  
 SOFTWARE: Patent Ver. 3.2  
 SEQ ID NO 330  
 LENGTH: 1083  
 TYPE: PRT  
 ORGANISM: Mycobacterium leprae  
 US-10-468-356-330

Query Match 71.2%; Score 37; DB 17; Length 1083;  
 Best Local Similarity 77.8%; Pred. No. 6.7e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ALPIEYGPL 9  
 DB 753 ALPSNYGPL 761

RESULT 15  
 US-10-424-599-21353

Tue Nov 30 08:49:27 2004

us-10-008-377a-6.open.rapb

Page 5

```
; Sequence 213533, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5323) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 213533
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34845C.1.pep
US-10-424-599-213533

Query Match          69.2%; Score 36; DB 15; Length 287;
Best Local Similarity 70.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CY      1 ALPEYGPV 10
      |||||
Db      56 ALPEYDPV 65

Search completed: November 30, 2004, 08:06:36
Job time : 377.923 secs
```

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 / Search time 25.1282 Seconds  
(without alignments)  
26.392 Million cell updates/sec

Title: US-10-008-377a-6

Sequence: 1 ALPIEXGPLV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:  
1: /cgn2\_6/prodata/1/aa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/aa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/aa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/aa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	525	1 US-08-160-861-4	Sequence 4, Appl
2	52	100.0	572	1 US-08-160-861-3	Sequence 3, Appl
3	52	100.0	680	1 US-08-542-363-4	Sequence 4, Appl
4	52	100.0	680	3 US-09-100-089-4	Sequence 4, Appl
5	52	100.0	680	4 US-09-670-827-4	Sequence 4, Appl
6	52	100.0	680	4 US-09-827-949-4	Sequence 4, Appl
7	52	100.0	1620	1 US-08-542-363-2	Sequence 2, Appl
8	52	100.0	1620	3 US-09-100-089-2	Sequence 2, Appl
9	52	100.0	1620	4 US-09-670-827-2	Sequence 2, Appl
10	52	100.0	1620	4 US-09-827-949-2	Sequence 2, Appl
11	38	73.1	207	4 US-09-543-681A-6664	Sequence 6664, Ap
12	38	73.1	409	4 US-09-252-991A-18004	Sequence 18004, A
13	37	71.2	206	4 US-09-562-737-55	Sequence 55, Appl
14	36	69.2	243	4 US-09-252-991A-59268	Sequence 29268, A
15	36	69.2	393	4 US-09-270-767-38581	Sequence 38581, A
16	36	69.2	393	4 US-09-270-767-53798	Sequence 53798, A
17	35	67.3	51	1 US-08-437-841-2	Sequence 2, Appl
18	35	67.3	51	1 US-08-286-521-2	Sequence 2, Appl
19	35	67.3	51	1 US-08-436-175-2	Sequence 2, Appl
20	35	67.3	51	3 US-08-943-682-2	Sequence 2, Appl
21	35	67.3	51	3 US-09-013-896A-7	Sequence 7, Appl
22	35	67.3	51	4 US-09-827-948-7	Sequence 7, Appl
23	35	67.3	51	4 US-09-741-105-2	Sequence 2, Appl
24	35	67.3	51	5 PCT-US95-09464-2	Sequence 2, Appl
25	35	67.3	55	1 US-08-147-710-15	Sequence 15, Appl
26	35	67.3	55	1 US-08-458-090-15	Sequence 15, Appl
27	35	67.3	55	2 US-08-457-887-15	Sequence 15, Appl

28	35	67.3	55	4 US-09-904-621-15	Sequence 15, Appl
29	35	67.3	57	4 US-09-144-428-21	Sequence 21, Appl
30	35	67.3	58	1 US-08-358-160-80	Sequence 80, Appl
31	35	67.3	58	1 US-08-676-125A-39	Sequence 39, Appl
32	35	67.3	58	2 US-09-136-012A-39	Sequence 39, Appl
33	35	67.3	58	3 US-08-676-124-70	Sequence 70, Appl
34	35	67.3	58	3 US-09-414-878-70	Sequence 70, Appl
35	35	67.3	58	3 US-09-240-136-70	Sequence 70, Appl
36	35	67.3	58	4 US-09-638-770A-70	Sequence 70, Appl
37	35	67.3	124	4 US-09-800-170-10	Sequence 10, Appl
38	35	67.3	213	2 US-08-796-850-2	Sequence 2, Appl
39	35	67.3	213	4 US-08-147-710-2	Sequence 2, Appl
40	35	67.3	235	1 US-08-457-887-2	Sequence 2, Appl
41	35	67.3	235	1 US-08-458-090-2	Sequence 2, Appl
42	35	67.3	235	2 US-08-457-887-2	Sequence 2, Appl
43	35	67.3	235	3 US-08-817-145-3	Sequence 3, Appl
44	35	67.3	235	4 US-09-702-705-332	Sequence 332, App
45	35	67.3	235	4 US-09-736-457-332	Sequence 332, App

## ALIGNMENTS

RESULT 1  
US-08-160-861-4  
Sequence 4, Application US/08160861  
Patent No. 5523295  
GENERAL INFORMATION:  
APPLICANT: MORRIS, STEPHAN W  
TITLE OF INVENTION: NOVEL FUSION NUCLEIC ACID SEQUENCES AND  
TITLE OF INVENTION: FUSION PROTEINS PRESENT IN HUMAN T(2:5) LYMPHOMA, METHODS  
OF DETECTION AND USES THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX  
STREET: 1100 NEW YORK AVE NW SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/160,861  
FILING DATE: 02-DEC-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MILLMAN, ROBERT A  
REGISTRATION NUMBER: 36217  
REFERENCE/DOCKET NUMBER: 0656.0400000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2678  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 525 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-160-861-4

Query Match 100.0%; Score 52; DB 1; Length 525;  
Best Local Similarity 100.0%; Pred. No. 0.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEXGPLV 10  
|||  
DB 456 ALPIEXGPLV 465

RESULT 2  
US-08-160-861-3

```
Sequence 3, Application US/08160861
Patent No. 5529295
GENERAL INFORMATION:
APPLICANT: MORRIS, STEPHAN W
TITLE OF INVENTION: NOVEL FUSION NUCLEIC ACID SEQUENCES AND
TITLE OF INVENTION: FUSION PROTEINS PRESENT IN HUMAN T(2:5) LYMPHOMA, METHODS
OF DETECTION AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVE NW SUITE 600
CITY: WASHINGTON
STATE: D.C.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,861
FILING DATE: 02-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A
REGISTRATION NUMBER: 36217
REFERENCE/DOCKET NUMBER: 0656.0400000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2678
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 572 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-160-861-3

Query Match
Best Local Similarity 100.0%; Score 52; DB 1; Length 572;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10
DB 503 ALPIEYGPLV 512

RESULT 3
US-08-542-363-4
Sequence 4, Application US/08542363
Patent No. 5770421
GENERAL INFORMATION:
APPLICANT: MORRIS, STEPHAN W.
APPLICANT: LOOK, A. THOMAS
TITLE OF INVENTION: ALK PROTEIN TYROSINE KINASE/RECEPTOR AND
TITLE OF INVENTION: LIGANDS THEREOF
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSER: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/542,363
FILING DATE: 12-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
```

```
NAME: FOX, SAMUEL L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0400001/SLE/GKT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 680 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-542-363-4

Query Match
Best Local Similarity 100.0%; Score 52; DB 1; Length 680;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10
DB 456 ALPIEYGPLV 465

RESULT 4
US-09-100-089-4
Sequence 4, Application US/09100089
Patent No. 6174674
GENERAL INFORMATION:
APPLICANT: MORRIS, STEPHAN W.
APPLICANT: LOOK, A. THOMAS
TITLE OF INVENTION: ALK PROTEIN TYROSINE KINASE/RECEPTOR AND
TITLE OF INVENTION: LIGANDS THEREOF
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSER: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,089
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/160,861
FILING DATE: 03-DEC-1993
APPLICATION NUMBER: US 08/542,363
FILING DATE: 12-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: FOX, SAMUEL L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0400002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 680 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-100-089-4

Query Match
Best Local Similarity 100.0%; Score 52; DB 3; Length 680;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 ALPIEXGPLV 10  
DB 456 ALPIEXGPLV 465

## RESULT 5

US-09-670-827-4

Sequence 4, Application US/09670827

Patent No. 6451997

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

Look, A. Thomas

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and

Ligands Thereof

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN &amp; FOX P.L.L.C.

STREET: 1100 New York Avenue, N.W., Suite 600

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/670,827

FILING DATE: 28-Sep-2000

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/160,861

FILING DATE: 03-Dec-1993

APPLICATION NUMBER: US 08/542,363

FILING DATE: 12-Oct-1995

APPLICATION NUMBER: US 09/100,089

FILING DATE: 19-Jun-1998

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 680 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-670-827-4

Query Match 100.0%; Score 52; DB 4; Length 680;

Best Local Similarity 100.0%; Pred. No. 0.27;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEXGPLV 10  
DB 456 ALPIEXGPLV 465

## RESULT 6

US-09-827-949-4

Sequence 4, Application US/09827949

Patent No. 6636548

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

Look, A. Thomas

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof

FILE REFERENCE: 0656.0400004

QY 1 ALPIEXGPLV 10  
DB 456 ALPIEXGPLV 465

RESULT 7

US-08-542-363-2

Sequence 2, Application US/08542363

Patent No. 5770421

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

Look, A. Thomas

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and

Ligands Thereof

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 New York Avenue, N.W., Suite 600

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/542,363

FILING DATE: 12-Oct-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400001/SLF/GKT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1620 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-542-363-2

Query Match 100.0%; Score 52; DB 4; Length 680;

Best Local Similarity 100.0%; Pred. No. 0.27;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 52; DB 1; Length 1620;

Best Local Similarity 100.0%; Pred. No. 0.67;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 ALPIEYGPLV 10
          |||||
Db      1396 ALPIEYGPLV 1405

```

RESULT 8  
US-09-100-089-2

```

: GENERAL INFORMATION:
:
: APPLICANT: Morris, Stephan W.
:
: APPLICANT: Look, A. Thomas
:
: TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and
:
: TITLE OF INVENTION: Ligands Thereof
:
: NUMBER OF SEQUENCES: 43
:
: COMPLETANUMBER NUMBER:

```

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: Patentn Release #1., Version #1.30
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/09/100,089
8  FILING DATE: Herewith

```

Query Match	100.0%	Score 52;	DB 3;	Length 1620;
Best Local Similarity	100.0%	Pred. No. 0.67;		
Matches	10;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0
QY	1	ALPRTYGPV	10	
Db	1336	ALPRTYGPV	1405	

RESULT 9  
 US-09-670-827-2  
 ; Sequence 2, Application US/09670827  
 ; Patent No. 645197  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Morris, Stephan W.  
 ; Inventor: look, J. Thomas  
 ; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
 ; Ligands Thereof  
 ; NUMBER OF SEQUENCES: 43  
 ; CORRESPONDENCE ADDRESS:  
 ;

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC

```

1      COMPUTER READABLE FORM:
2      MEDIUM TYPE: Floppy disk
3      COMPUTER: IBM PC compatible
4      OPERATING SYSTEM: PC-DOS/MS-DOS
5      SOFTWARE: PatentIn Release #1.0, Version #1.30
6      CURRENT APPLICATION DATA:
7      APPLICATION NUMBER: US/09/670, 827
8      FILING DATE: 28-Sep-2000
9      CLASSIFICATION: unknown
10     PRIOR APPLICATION DATA: US 08/160,964

```

Query Match	100.0%;	Score 52;	DB 4;	Length 1620;
Best Local Similarity	100.0%;	Pred. No. 0.67;		
Matches	10;	Conservative	0;	Indels 0;
		Mismatches	0;	Gaps 0;

```

RESULT 10
US-09-827-949-2
; Sequence 2, Application US/09827949
; Patent No. 6696548
; GENERAL INFORMATION:
; APPLICANT: Morris, Stephan W.
; APPLICANT: LOOK, A. Thomas
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof
; FILE REFERENCE: 0656.0400004
; CURRENT APPLICATION NUMBER: US/09/827,949
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 09/670,827
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/100,089
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: US 08/542,363
; PRIOR FILING DATE: 1995-10-12
; PRIOR APPLICATION NUMBER: US 08/160,861
; PRIOR FILING DATE: 1993-12-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 1620
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-09-827-949-2

Query Match 100.0%; Score 52; DB 4; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALPIEYGPLV 10  
Db 1396 ALPIEYGPLV 1405

RESULT 11  
US-09-543-681A-6664  
Sequence 6664, Application US/09543681A  
Patent No. 6605708  
GENERAL INFORMATION:  
APPLICANT: GARY BRETTON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543,681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 6664  
LENGTH: 207  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-09-543-681A-6664

Query Match 73.1%; Score 38; DB 4; Length 207;  
Best Local Similarity 60.0%; Pred. No. 24;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALPIEYGPLV 10  
Db 110 ALDMEYGPIL 119

RESULT 12  
US-09-252-991A-18004  
Sequence 18004, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: MARC J. RUBENFELD et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 18004  
LENGTH: 409  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18004

Query Match 73.1%; Score 38; DB 4; Length 409;  
Best Local Similarity 88.9%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALPIEYGPLV 9  
Db 286 ALPIEYGPLV 294

RESULT 13  
US-09-562-737-55

Sequence 55, Application US/09562737

Patent No. 6428967  
GENERAL INFORMATION:  
APPLICANT: Herz, Joachim  
APPLICANT: Gotthardt, Michael  
TITLE OF INVENTION: LDL Receptor Signaling Pathways  
FILE REFERENCE: UTSW0708  
CURRENT APPLICATION NUMBER: US/09/562,737  
CURRENT FILING DATE: 2000-05-01  
NUMBER OF SEQ ID NOS: 132  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 55  
LENGTH: 206  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-562-737-55

Query Match 71.2%; Score 37; DB 4; Length 206;  
Best Local Similarity 55.6%; Pred. No. 36;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LPVQWGPV 10  
Db 161 LPVQWGPV 169

RESULT 14  
US-09-252-991A-29268  
Sequence 29268, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: MARC J. RUBENFELD et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 29268  
LENGTH: 243  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29268

Query Match 69.2%; Score 36; DB 4; Length 243;  
Best Local Similarity 75.0%; Pred. No. 65;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALPIEYGPLV 8  
Db 206 ALPIEYGPLV 213

RESULT 15  
US-09-270-767-38581  
Sequence 38581, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 38581

LENGTH: 393  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
FEATURE:  
OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-38581

Query Match 69.2%; Score 36; DB 4; Length 393;  
Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 3 PIRXGPI 9  
Db 107 PIRGPI 113

Search completed: November 30, 2004, 07:19:16  
Job time : 26.1282 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 18 Seconds  
(without alignments)  
48.108 Million cell updates/sec

Title: US-10-008-377a-7  
Perfect score: 42  
Sequence: 1 SLTNMKEV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	42	100.0	680 4	138491
2	33	78.6	206 2	122345
3	32	76.2	341 2	A11186
4	31	73.8	510 2	A11186
5	31	73.8	510 2	AD1554
6	31	73.8	544 2	AH1544
7	31	73.8	583 2	S18642
8	31	73.8	599 2	F82291
9	30	71.4	145 2	AF1259
10	30	71.4	232 2	C95196
11	30	71.4	232 2	H98062
12	30	71.4	232 2	C98021
13	30	71.4	233 2	C95154
14	30	71.4	285 2	A86732
15	30	71.4	291 2	S61498
16	30	71.4	340 2	S20879
17	30	71.4	347 2	AF2645
18	30	71.4	347 2	F97427
19	30	71.4	550 1	FGRTA
20	30	71.4	551 2	S52287
21	30	71.4	608 2	A72213
22	30	71.4	1734 2	A54602
23	29	69.0	184 2	S55435
24	29	69.0	198 2	S86771
25	29	69.0	215 2	E75058
26	29	69.0	221 2	A11660
27	29	69.0	222 2	D65882
28	29	69.0	222 2	E72041
29	29	69.0	222 2	AC1285

30	29	69.0	243 2	D82681	protein transferas
31	29	69.0	359 2	H90301	hypothetical prote
32	29	69.0	363 2	AB1537	hypothetical prote
33	29	69.0	395 2	T40102	conserved hypothet
34	29	69.0	398 2	F64456	hypothetical prote
35	29	69.0	418 2	A53120	intracellular coag
36	29	69.0	482 2	T39455	adenylosuccinate 1
37	29	69.0	482 2	S51377	probable membrane
38	29	69.0	521 2	C28529	nicotinic acetylch
39	29	69.0	611 2	T45493	glutamine-fructose
40	29	69.0	625 2	T25373	hypothetical prote
41	29	69.0	677 2	S15667	transcription fact
42	29	69.0	701 2	S17196	transcription fact
43	29	69.0	727 2	S18193	autoantigen NOR-90
44	29	69.0	727 2	JC5113	ribosomal transcri
45	29	69.0	727 2	B40439	UBF transcription

## ALIGNMENTS

RESULT 1  
138491  
nucleophosmin/anaplastic lymphoma kinase mutant fusion protein - human  
C:Species: Homo sapiens (man)  
C:Date: 15-Feb-1996 #sequence\_revision 15-Feb-1996 #text\_change 20-Apr-2000  
C:Accession: 138491  
R:Morris, S.W.; Kirstein, M.N.; Valentine, M.B.; Dittmer, K.G.; Shapiro, D.N.; Saltman, Science 263, 1281-1284, 1994  
A>Title: Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin  
A:Reference number: A53182; MUID:94167588; PMID:8122112  
A:Accession: 138491  
A:Molecule type: mRNA  
A:Residues: 1-680 <NOR>  
A:Cross-references: EMBL:U04946; NID:G609341; PIDN:AA58698.1; PID:G609342  
C:Comment: This sequence is the chimeric product of a translocation mutation.  
C:Genetics:  
A:Gene: NPM1/ALK  
A:Map position: 5/2p23-2p23  
C:Keywords: fusion protein

Query Match 100.0%; Score 42; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTNMKEV 9  
DB 621 SLTNMKEV 629

RESULT 2  
T2345  
hypothetical protein F478.8 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: 122345  
R:Berkis, M.; McMurtry, A.  
A:Submitted to the EMBL Data Library, July 1996  
A:Reference number: Z19551  
A:Accession: T2345  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-206 <NLS>  
A:Cross-references: UNIPROT:Q20521; EMBL:Z77662; PIDN:CA801196.1; GSERDB:GN00023; CESP:F  
A:Experimental source: clone F4788  
C:Genetics:  
A:Gene: CESP:F4788.8  
A:Map position: 5  
A:introns: 71/1, 171/3

Query Match 78.6%; Score 33; DB 2; Length 206;  
Best Local Similarity 75.0%; Pred. No. 10;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLTANKKEV 8  
 Db 131 SLTANKMDL 138

## RESULT 3

transport proteins homolog lmo0897 [imported] - *Listeria monocytogenes* (strain EGD-e)  
 C/Species: *Listeria monocytogenes*  
 C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: A11186  
 R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Biocker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A/Authors: Kretz, U.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A./Title: Comparative genomics of *Listeria* species.  
 A/Reference number: AB1077; MUID:21537279; PMID:11679669  
 A/Accession: A11186  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-541 <G>  
 A/Cross-references: UNIPROT:Q8Y8K4; GB:NC\_003210; PIDN:CAC98975.1; PID:G16410300; GSPDB: A/Experimental source: strain EGD-e  
 C/Genetics:  
 A/Gene: lmo0897  
 C/Superfamily: integral membrane protein HP0228

Query Match 76.2%; Score 32; DB 2; Length 541;  
 Best Local Similarity 66.7%; Pred. No. 50;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SLTANKKEV 9  
 Db 154 SFVANKKEI 162

## RESULT 4

D-alanine-activating enzyme (dae), D-alanine-D-alanyl carrier protein ligase (dcl) [impoc  
 C/Species: *Listeria monocytogenes*  
 C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: A11196  
 R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Biocker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A/Authors: Kretz, U.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A./Title: Comparative genomics of *Listeria* species.  
 A/Reference number: AB1077; MUID:21537279; PMID:11679669  
 A/Accession: A11196  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-510 <G>  
 A/Cross-references: UNIPROT:Q8Y8D4; GB:NC\_003210; PIDN:CAC99052.1; PID:G16410376; GSPDB: A/Experimental source: strain EGD-e  
 C/Genetics:  
 A/Gene: dclA  
 C/Superfamily: acetate-CoA ligase; acetate-CoA ligase homology

Query Match 73.8%; Score 31; DB 2; Length 510;  
 Best Local Similarity 55.6%; Pred. No. 78;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLTANKKEV 9  
 Db 224 TITANKMDL 232

## RESULT 5

AD1554  
 D-alanine-activating enzyme (dae), D-alanine-D-alanyl carrier protein ligase (dcl) [impoc  
 C/Species: *Listeria innocua*  
 C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: AD1554  
 R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Biocker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A/Authors: Kretz, U.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A./Title: Comparative genomics of *Listeria* species.  
 A/Reference number: AB1077; MUID:21537279; PMID:11679669  
 A/Accession: AD1554  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-510 <G>  
 A/Cross-references: UNIPROT:Q92D47; GB:AL592022; PIDN:CAC96204.1; PID:G16413342; GSPDB: A/Experimental source: strain C1p11262  
 C/Genetics:  
 A/Gene: dclA

Query Match 73.8%; Score 31; DB 2; Length 510;  
 Best Local Similarity 55.6%; Pred. No. 78;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLTANKKEV 9  
 Db 224 TITANKMDL 232

## RESULT 6

transport protein homolog lln0896 [imported] - *Listeria innocua* (strain C1p11262)  
 C/Species: *Listeria innocua*  
 C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: A11544  
 R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Biocker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A/Authors: Kretz, U.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A./Title: Comparative genomics of *Listeria* species.  
 A/Reference number: AB1077; MUID:21537279; PMID:11679669  
 A/Accession: A11544  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-544 <G>  
 A/Cross-references: UNIPROT:Q92DB9; GB:AL592022; PIDN:CAC96128.1; PID:G16413346; GSPDB: A/Experimental source: strain C1p11262  
 C/Genetics:  
 A/Gene: lln0896  
 C/Superfamily: integral membrane protein HP0228

Query Match 73.8%; Score 31; DB 2; Length 544;  
 Best Local Similarity 66.7%; Pred. No. 84;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SLTANKKEV 9  
 Db 154 SFVANKKEI 162

## RESULT 7

chaperonin 60 alpha chain precursor, chloroplast - rae  
 C/Species: *Brassica napus* (rape)  
 C/Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
 C/Accession: S38642  
 R/Cole, K.P.; Blakeley, S.D.; Dennis, D.T.  
 submitted to the EMBL Data Library, November 1993  
 A/Description: Isolation of a full length cDNA encoding *Brassica napus* plastid chaperon





\_\_\_\_\_

[REDACTED]

C98021  
N-acetylmannosamine-6-P epimerase [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: C98021  
R:Host: J.A.; Albom Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.C.; Lu, J.; Matsushima, P.; McAhren, S.; M  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: C98021  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-232 <KUS>  
A:Cross-references: UNIPROT:Q8DPF0; GB:AE007317; PDB:1AK99999.1; PDB:1J5458829; GSPDB:C  
C:Genetics:  
A:Gene: nanB

Query Match 71.4%; Score 30; DB 2; Length 232;  
Best Local Similarity 75.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LTANKREV 9  
DB 86 ITANKREV 93

RESULT 13  
C95154  
N-acetylmannosamine-6-P epimerase, probable [imported] - Streptococcus pneumoniae (strain  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: C95154  
R:Retelien, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Kadane, D.; Holtapple,  
neon, T.; Hickey, E.K.; Holt, I.B.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: C95154  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-233 <KUS>  
A:Cross-references: UNIPROT:Q97095; GB:AE005672; PDB:1AK75428.1; PDB:1J4972812; GSPDB:C  
A:Experimental source: strain ITGR4  
C:Genetics:  
A:Gene: SPI330

Query Match 71.4%; Score 30; DB 2; Length 233;  
Best Local Similarity 75.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LTANKREV 9  
DB 87 ITANKREV 94

RESULT 14  
A86732  
geranyltransferase (EC 2.5.1.10) [imported] - Lactococcus lactis subsp. lactis (str  
N:Alternate names: farnesyl diphosphate synthase  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: A86732  
R:Boletín, A.; Winkler, P.; Mauger, S.; Jallón, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A:Reference number: A86625; MUID:21235186; PMID:11337471  
A:Accession: A86732  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-285 <STO>  
A:Cross-references: UNIPROT:Q9CH81; GB:AE005176; PDB:1G1272785; PDB:1AK04955.1; GSPDB:C  
A:Experimental source: strain ILL403  
C:Genetics:  
A:Gene: ispA  
C:Keywords: transferase

Query Match 71.4%; Score 30; DB 2; Length 285;  
Best Local Similarity 66.7%; Pred. No. 68;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LTANKREV 9  
DB 252 LTANKREV 260

RESULT 15  
S61498  
chemotactic methyltransferase homolog pilK - Pseudomonas aeruginosa  
C:Species: Pseudomonas aeruginosa  
C:Date: 27-Apr-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S61498; D83594  
R:Darzi, A.  
Mol. Microbiol. 15, 703-717, 1995  
A:Title: The Pseudomonas aeruginosa pilK gene encodes a chemotactic methyltransferase ('  
A:Reference number: S61498; MUID:95302983; PMID:7783642  
A:Accession: S61498  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-291 <DAR>  
A:Cross-references: UNIPROT:Q51346; EMBL:U11382; NID:9520528; PDB:1AA85790.1; PDB:9520  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; Lim,  
Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
A:Reference number: A82950; MUID:20437337; PMID:10964043  
A:Accession: D83594  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-291 <STO>  
A:Cross-references: GB:AE004478; GB:AE004091; NID:99946261; PDB:1AAG3801.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: pilK; PA0412  
A:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransfer  
F,13-274/DNA: protein-glutamate O-methyltransferase homolog <PKM>

Query Match 71.4%; Score 30; DB 2; Length 291;  
Best Local Similarity 66.7%; Pred. No. 70;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 LTANKREV 9  
DB 46 LTANKREV 54

Search completed: November 30, 2004, 07:41:26  
Job time: 20 secs

Tue Nov 30 08:49:33 2004

us-10-008-377a-7.open.rup

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 : Search time 68.7692 Seconds  
(without alignments)  
75.301 Million cell updates/sec

Title: US-10-008-377a-7  
Perfect score: 42  
Sequence: 1 SLTANKKEV 9

Scoring table: ELOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	803	2	Q8TDJ5
2	42	100.0	1620	2	Q8TDJ5
3	35	83.3	1194	2	Q8R6P0
4	34	81.0	256	2	Q8A0P1
5	33	78.6	206	2	Q20521
6	33	78.6	660	2	Q17248
7	33	78.6	807	2	Q81B74
8	32	76.2	320	2	Q6TV42
9	32	76.2	320	2	AAR39404
10	32	76.2	510	2	Q8W2N4
11	32	76.2	541	2	Q8Y8K4
12	32	76.2	684	2	Q6ZKX5
13	32	76.2	684	2	Q6ZKX5
14	31	73.8	124	2	Q7ACFE
15	31	73.8	124	2	Q7ACFE
16	31	73.8	170	2	AAR35086
17	31	73.8	170	2	Q6K262
18	31	73.8	238	2	Q6T2V4
19	31	73.8	250	2	AAR3901
20	31	73.8	254	2	Q8N4X4
21	31	73.8	254	2	Q8R6U6
22	31	73.8	309	2	Q8GY34
23	31	73.8	367	2	Q839S1
24	31	73.8	510	1	DLTA_LISTN
25	31	73.8	510	1	DLTA_LISTN
26	31	73.8	510	1	Q72HJ2
27	31	73.8	510	2	AAT03772
28	31	73.8	544	2	Q92DB9
29	31	73.8	583	1	RUPA_BRANA
30	31	73.8	583	1	Q6BVC7
31	31	73.8	598	2	Q73HR8
31	31	73.8	598	2	AAS14195

32	31	73.8	599	2	Q87JH6	Q87JH6 vibrio para
33	31	73.8	599	2	Q87JH6	Q87JH6 vibrio para
34	31	73.8	764	2	Q76130	Q76130 neurospora
35	31	73.8	862	1	DR33_ARATH	DR33_ARATH arabidopsis
36	31	73.8	1113	2	Q6CMF0	Q6CMF0 kiuyveromyc
37	30	71.4	110	2	Q6ELI3	Q6ELI3 biophthalari
38	30	71.4	133	2	Q9N3J4	Q9N3J4 caenorhabdi
39	30	71.4	145	2	Q8Y743	Q8Y743 listeria mo
40	30	71.4	145	2	Q71ZU2	Q71ZU2 listeria mo
41	30	71.4	145	2	AAT04272	AAT04272 listeria
42	30	71.4	179	1	ATPD_STRAEP	ATPD_STRAEP streptococ
43	30	71.4	212	2	Q6NN76	Q6NN76 dirosophila
44	30	71.4	212	2	AAR6209	AAR6209 dirosophila
45	30	71.4	232	1	NAEL_STRRE	Q8P60 streptococ

#### ALIGNMENTS

RESULT 1	ID	Q8TDJ5	PRELIMINARY;	PRT;	803 AA.
AC	Q8TDJ5	01-JUN-2002 (TREMURel. 21, Created)			
DT	01-JUN-2002 (TREMURel. 21, Last sequence update)				
DT	01-MAR-2004 (TREMURel. 26, Last annotation update)				
DE	TRK-fused gene/anaplastic large cell lymphoma kinase extra long				
DE	form.				
GN	Name=TRF/ALK fusion;				
OC	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid=9606;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21940189; PubMed=11943732;				
RA	Hernandez L., Bea S., Bellocillo B., Pinyol M., Falini B., Ott G.,				
RA	Pulford K., Rosenwald A., Morris S.W., Fernandez A., Santos E.,				
RA	Campo E.;				
RT	"Diversity of genomic breakpoints in TFG-ALK translocations in				
RT	anaplastic large cell lymphomas: identification of a new TFG-ALK (XL)				
RT	chimeric gene with transforming activity."				
RL	Am. J. Pathol. 160:1487-1494(2002).				
DR	EMBL: AF390893; AAM17922.1; -				
DR	HSP: Q62839; ILUP				
DR	GO: GO:0005524; F:ATP binding; IEA.				
DR	GO: GO:0004723; F:Protein-tyrosine kinase activity; IEA.				
DR	GO: GO:0016740; F:Transferase activity; IEA.				
DR	GO: GO:0006468; P:Protein amino acid phosphorylation; IEA.				
DR	InterPro: IPR011009; Kinase_Like.				
DR	InterPro: IPR000270; OPR_PBI.				
DR	InterPro: IPR000719; Prot_kinase.				
DR	InterPro: IPR002011; ReceptTyknsII.				
DR	InterPro: IPR001245; Tyr_kinase.				
DR	InterPro: IPR008266; Tyr_kinase_AS.				
DR	Pfam: PF00564; PBI.1				
DR	Pfam: PF00069; PKINASE.1				
DR	PRINTS: PR00109; TYRKINASE.				
DR	PRODOM: PD000001; Prot_kinase; 1.				
DR	SMART: SM00666; PBI.1				
DR	SMART: SM00219; TYKc.1				
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.				
DR	PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.				
DR	PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.				
DR	PROSITE: PS00239; RECEPTOR_TYR_KIN_II; UNKNOWN_1.				
KW	ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.				
SEQ	SEQUENCE 803 AA; 88670 MW; 89082645956598F CRC64;				

Query Match 100.0%; Score 42; DB 2; Length 803;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLTANKKEV 9

DB 744 SLTANKEV 752

RESULT 2  
ID ALK\_HUMAN STANDARD; PRT; 1620 AA.  
AC Q9UM73; O9Y4K6;  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE ALK tyrosine kinase receptor precursor (EC 2.7.1.112) (Anaplastic  
DE lymphoma kinase) (CD246 antigen).  
GN Name:ALK;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND  
RP GLYCOSYLATION.  
RX MEDLINE=97316779; PubMed=9174053; DOI=10.1038/sj.onc.1201062;  
RA Morris S.W., Naeye C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,  
RA Witte D.P.;  
RT "ALK, the chromosome 2 gene locus altered by the t(2;5) in non-  
RT Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase  
RT that is highly related to leukocyte tyrosine kinase (LTK).";  
RL Oncogene 14:2175-2188(1997).  
RN [2]  
RP ERRATUM.  
RA Morris S.W., Naeye C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,  
RA Witte D.P.;  
RL Oncogene 15:2883-2883(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97178863; PubMed=9053841;  
RA Iwabara T., Fujimoto T., Wen D., Cupples R., Bucay N., Araiawa T.,  
RA Mori S., Ratzkin B., Yamamoto T.;  
RT "Molecular characterization of ALK, a receptor tyrosine kinase  
RT expressed specifically in the nervous system.";  
RL Oncogene 14:439-449(1997).  
RN [4]  
RP PARTIAL SEQUENCE FROM N.A., AND CHROMOSOMAL TRANSLOCATION.  
RX MEDLINE=94167588; PubMed=8122112;  
RA Morris S.W., Kirstein M.N., Valentine M.B., Dittmer K.G.,  
RA Shapiro D.N., Saltman D.L., Look A.T.;  
RT "Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in  
RT non-Hodgkin's lymphoma.";  
RL Science 263:1281-1284(1994).  
CC -1- FUNCTION: Orphan receptor with a tyrosine-protein kinase activity.  
CC Appears to play an important role in the normal development and  
CC function of the nervous system.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: Expressed in brain and CNS. Also expressed in  
CC the small intestine and testis, but not in normal lymphoid cells.  
CC -1- PTM: N-glycosylated.  
CC -1- DISEASE: A form of non-Hodgkin's lymphoma is characterized by a  
CC chromosomal translocation t(2;5)(p23;q35) that involves NPM1 and  
CC ALK.  
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin  
CC receptor subfamily.  
CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.  
CC -1- SIMILARITY: Contains 2 WML domains.  
CC -1- DATABASE: NMRAtlas Genet. Cytogenet. Oncol. Haematol.;  
CC WWW="http://www.inbiochem.fr/services/cytocancer/genes/ALK.html".  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb.ch](mailto:license@isb.ch)).  
CC -----  
CC EMBL; U62540; AAB71619.1; -  
CC EMBL; U66559; AAC51104.1; -  
CC HSP; Q62838; IJUF.  
CC Genew; HGNC:427; ALK.  
CC MIM; 105590;  
CC GO; GO:0016021; C: integral to membrane; NAS.  
CC GO; GO:007399; P: neurogenesis; NAS.  
CC InterPro; IPR011009; Kinase-like.  
CC InterPro; IPR002172; LDL\_receptor\_A.  
CC InterPro; IPR000988; MAM.  
CC InterPro; IPR000719; Prot. kinase.  
CC InterPro; IPR002011; Recept.tyrosin.  
CC InterPro; IPR001245; Tyr. kinase.  
CC InterPro; IPR008266; Tyr\_kinase\_AS.  
CC Pfam; PF00692; MAM.1.  
CC Pfam; PF00063; Kinase.1.  
CC PRINTS; PR00109; TYRKINASE.  
CC ProDom; PD000001; Prot\_kinase.1.  
CC SMART; SM00192; Ldla; 1.  
CC SMART; SM00219; Tyrc; 1.  
CC PROSITE; PS01209; LDLRA\_1; FALSE NEG.  
CC PROSITE; PS00068; LDLRA\_2; FALSE NEG.  
CC PROSITE; PS00740; MAM\_1; FALSE NEG.  
CC PROSITE; PS00060; MAM\_2; 2.  
CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP.1.  
CC PROSITE; PS00111; PROTEIN\_KINASE\_DOM.1.  
CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR.1.  
CC PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II.1.  
CC ATP-binding; Chromosomal translocation; Glycoprotein; Phosphorylation;  
CC Proto-oncogene; Receptor; Repeat; Signal; Transferase; Transmembrane;  
CC Tyrosine-protein kinase.  
CC KW SIGNAL.  
CC FT CHAIN 1 18  
CC FT 19 1620  
CC FT DOMAIN 19 1038  
CC FT TRANSFAM 1039 1055  
CC FT DOMAIN 1060 1620  
CC FT DOMAIN 264 427  
CC FT DOMAIN 437 478  
CC FT DOMAIN 478 636  
CC FT DOMAIN 1116 1392  
CC FT DOMAIN 816 940  
CC FT NP\_BIND 1122 1130  
CC FT BINDING 1150 1150  
CC FT ACT\_SITE 1249 1249  
CC FT MOD\_RES 1282 1282  
CC FT CARBOHYD 169 169  
CC FT CARBOHYD 244 244  
CC FT CARBOHYD 285 285  
CC FT CARBOHYD 324 324  
CC FT CARBOHYD 411 411  
CC FT CARBOHYD 424 424  
CC FT CARBOHYD 445 445  
CC FT CARBOHYD 563 563  
CC FT CARBOHYD 571 571  
CC FT CARBOHYD 627 627  
CC FT CARBOHYD 709 709  
CC FT CARBOHYD 808 808  
CC FT CARBOHYD 863 863  
CC FT CARBOHYD 864 864  
CC FT CARBOHYD 886 886  
CC FT CARBOHYD 986 986  
CC FT CONFLICT 36 36  
CC FT CONFLICT 1491 1491  
CC FT CONFLICT 1529 1529  
CC SQ SEQUENCE 1620 AA; 176417 MW; A626049824296181E CRC64;  
Query Match 100.0%; Score 42; DB 1; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 6.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 DB 1561 SLTANKEV 1569

## RESULT 3

Q9R6FO PRELIMINARY; PRT: 1194 AA.  
 AC Q9R6FO: 01-MAY-2000 (TREMUREL. 13, Created)  
 DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)  
 DT 01-JUN-2003 (TREMUREL. 24, Last annotation update)  
 DE Tior109 protein.  
 GN Name=Tior109;  
 OS Agrobacterium tumefaciens.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiales; Rhizobium/Agrobacterium group; Agrobacterium.  
 CX NCBI\_TaxID=558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RA Hattori Y., Suzuki K., Ohta N., Katoh A., Yoshida K.;  
 RT "Genome structure of pTi-SAKURA (II): Strategy for DNA sequencing of a  
 RT Japanese cherry-Ti plasmid.";  
 RL Nucleic Acids Symp. Ser. 37:159-160 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RA Ohta N., Suzuki K., Hattori Y., Uraji M., Katoh A., Yoshida K.;  
 RT "Genome structure of pTi-SAKURA (III): Characteristics of T-DNA.";  
 RL Nucleic Acids Symp. Ser. 39:185-186 (1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RA Uraji M., Suzuki K., Ohta N., Hattori Y., Katoh A., Yoshida K.;  
 RT "Genome structure of pTi-SAKURA (IV): Characteristics of tra region.";  
 RL Nucleic Acids Symp. Ser. 39:187-188 (1998).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RA Hattori Y., Suzuki K., Ohta N., Katoh A., Yoshida K.;  
 RT "Genome structure of pTi-SAKURA (V): Complete nucleotide sequence of  
 RT plasmid pTi-SAKURA's vir region in Agrobacterium tumefaciens.";  
 RL Nucleic Acids Symp. Ser. 39:265-266 (1998).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RX MEDLINE=20184752; PubMed=1072127;  
 RA Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,  
 RA Katoh A., Yoshida K.;  
 RT "Complete nucleotide sequence of a plant tumor-inducing Ti plasmid.";  
 RL Gene 242:331-336 (2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RX MEDLINE=98193120; PubMed=9524202;  
 RA Suzuki K., Ohta N., Hattori Y., Uraji M., Katoh A., Yoshida K.;  
 RT "Novel structural difference between nopaline- and octopine- type trb  
 RT gene: construction of genetic and physical map and sequencing of  
 RT trb/trai and rep gene clusters of a new Ti plasmid pTi-SAKURA.";  
 RL Biochim. Biophys. Acta 1396:1-7 (1998).  
 DR EMBL: AB016260; BAA8734.1;  
 DR GO: GO:0003921; Pnnid:sectional conjugation; IEA.  
 DR InterPro:IPR005053; MobX\_MobX.  
 DR Pfam: Pf03389; MobX\_MobX.1.  
 KW Plasmid.  
 SQ SEQUENCE 1194 AA; 133751 MW; C9B26E655349P96 CRC64;

Query Match 83.3%; Score 35; DB 2; Length 1194;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 DB 920 SLTANKEV 928

## RESULT 4

Q9A0P1 PRELIMINARY; PRT: 256 AA.  
 AC Q9A0P1: 01-JUN-2003 (TREMUREL. 24, Created)  
 DT 01-JUN-2003 (TREMUREL. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMUREL. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocusNames=BT3980;  
 OS Bacteroides thetaiotaomicron.  
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 OC Bacteroidales; Bacteroides.  
 CX NCBI\_TaxID=818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VPI-5482 / ATCC 29148;  
 RX MEDLINE=22550858; PubMed=1263928;  
 RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
 RA Chang H.C., Hooper L.V., Gordon J.I.;  
 RT "A genomic view of the human Bacteroides thetaiotaomicron symbiosis.";  
 RL Science 289:2074-2076 (2003).  
 DR EMBL: AB016943; AA079085.1;  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 256 AA; 30314 MW; 6EF4EA7645790EDC CRC64;

Query Match 81.0%; Score 34; DB 2; Length 256;  
 Best Local Similarity 77.8%; Pred. No. 49;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 DB 46 SLTANKEV 54

## RESULT 5

Q20521 PRELIMINARY; PRT: 206 AA.  
 AC Q20521: 01-NOV-1996 (TREMUREL. 01, Created)  
 DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)  
 DT 01-OCT-2003 (TREMUREL. 25, Last annotation update)  
 DE Hypothetical protein F47B8.8.  
 GN ORFNames=F47B8.8;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Petodermidae; Caenorhabditis.  
 CX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Belts M.;  
 RT Sumitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: Z77662; CAB01196.1;  
 DR PIR: T22345; T22345.  
 DR WormPep: F47B8.8; CE10646.  
 DR GO: GO:0004179; F:membrane alanyl aminopeptidase activity; IEA.  
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro:IPR001930; Peptidase\_M1.  
 DR Pfam: Pf01433; Peptidase\_M1.1.

Query Match 83.3%; Score 35; DB 2; Length 1194;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

KW Hypothetical protein.  
SQ SEQUENCE 206 AA; 23204 MW; A1952AA79BD86075 CRC64;

Query Match 78.6%; Score 33; DB 2; Length 206;  
Best Local Similarity 75.0%; Pred. No. 65;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKE 8  
DB 131 SLTANKE 138

## RESULT 6

ID 017248 PRELIMINARY; PRT; 660 AA.  
AC 017248;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)  
DE Angiotensin-converting enzyme-like protein precursor.  
GN Name=Ang2;  
OS Boophilus microplus (Cattle tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Boophilus.  
OX NCBI\_TaxID=6941;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Tissue-whole ticks;  
RA Whitefield P.L., Johnson M.C., Smith D.R.J., Snelson V.J., Clifton G.F.,  
RA Brown G.S., Cairns D., Foy A.B., Irving D.O.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U62809; AAB04998.1; -  
DR HSSP; Q10714; IJ36.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004246; F:peptidyl-dipeptidase A activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR01546; Peptidase\_M2.  
DR Pfam; PF01401; Peptidase\_M2; 1.  
DR PRINTS; PR00791; PEPDPTASEA.  
KW Signal.  
FT SIGNAL.  
FT CHAIN 1 29 Potential.  
SQ SEQUENCE 660 AA; 75257 MW; 6F164CF70C938E63 CRC64;  
Query Match 78.6%; Score 33; DB 2; Length 660;  
Best Local Similarity 87.5%; Pred. No. 2,4e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKEV 9  
DB 173 LTANKEV 180

## RESULT 7

ID 081874 PRELIMINARY; PRT; 807 AA.  
AC 081874;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, last annotation update)  
DE Hypothetical protein PF08\_0030.  
GN Name=PF08\_0030;  
OS Plasmodium falciparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
OX NCBI\_TaxID=36629;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Seegeer K., Murphy L., Harris D., Berriman M., Pain A., Hall N.,  
RA Quail M., Barrett B.;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL844507; CAD51133.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 807 AA; 95820 MW; 7CFB4D3A720625F2 CRC64;

Query Match 78.6%; Score 33; DB 2; Length 807;  
Best Local Similarity 87.5%; Pred. No. 2.9e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKE 8  
DB 764 SLTANKE 771

## RESULT 8

ID 06TV42 PRELIMINARY; PRT; 320 AA.  
AC 06TV42;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, last annotation update)  
DE Putative class II fructose-1,6-bisphosphatase / seduheptulose-1,7-bisphosphatase.  
GN Name=gipx;  
OS Bacillus methanolicus.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1471;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MG3;  
RA PubMed:14973041;  
RA Bratacset T., Jakobsen O.M., Flickinger M.C., Valla S.,  
RA Ellingsen T.E.;  
RL "Plasmid-dependent methylotrophy in thermotolerant Bacillus methanolicus."  
RT J. Bacteriol. 186:1229-1238(2004).  
DR EMBL; AY386314; AAR39404.1; -  
DR InterPro; IPR002453; Beta\_tubulin.  
DR InterPro; IPR004454; GIPX.  
DR Pfam; PF007014; GIPX; 1.  
DR Pfam; PF003320; FBpase\_gipx; 1.  
DR Pfam; PF007014; GIPX; 1.  
DR TIGRFAMs; TIGR00330; gipx; 1.  
DR PROSITE; PS00228; TUBULIN\_B\_AUTOREG; UNKNOWN\_1.  
KW Plasmid.  
SQ SEQUENCE 320 AA; 34557 MW; 816B36237154628C CRC64;  
Query Match 76.2%; Score 32; DB 2; Length 320;  
Best Local Similarity 77.8%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 146 SLTANKEV 154

## RESULT 9

ID AAR39404 PRELIMINARY; PRT; 320 AA.  
AC AAR39404;  
DT 02-MAR-2004 (TREMBlrel. 27, Created)  
DT 02-MAR-2004 (TREMBlrel. 27, last sequence update)  
DT 02-MAR-2004 (TREMBlrel. 27, last annotation update)  
DE Putative class II fructose-1,6-bisphosphatase / seduheptulose-1,7-bisphosphatase.  
GN GIPX.  
OS Bacillus methanolicus.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1471;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MG3;  
RA PubMed:14973041;  
RA Bratacset T., Jakobsen O.M., Flickinger M.C., Valla S.,  
RA Ellingsen T.E.;  
RL "Plasmid-dependent Methylotrophy in Thermotolerant Bacillus

```

RT methanolicus."
RL J. Bacteriol. 186:1229-1238(2004).
DR EMBL: AY386314; MAR39404.1; -.
KW plasmid.
SQ SEQUENCE 320 AA; 34557 MW; 815E36237154628C CRC64;

Query Match
Best Local Similarity 76.2%; Score 32; DB 2; Length 320;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 SLTANKEV 9
DB 146 SLTANKEV 154

RESULT 10
Q8W2N4 PRELIMINARY; PRT; 510 AA.
AC Q8W2N4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Cytochrome P450-dependent fatty acid hydroxylase.
GN Name=CYP9A6;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Lehouquin R., Kahn R., Benveniste I., Durst F.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL: AF092913; AAL54884.1; -.
DR GO: GO:0004497; F:monooxygenase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001128; Cytochrome_P450.
DR InterPro: IPR002401; EP4501.
DR Pfam: PF00067; P450_1.
DR PRINTS: PR00463; EP4501.
DR PRINTS: PR00385; P450.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 510 AA; 58633 MW; 67F0FD6472DAA182 CRC64;

Query Match
Best Local Similarity 76.2%; Score 32; DB 2; Length 510;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 LITANKEV 9
DB 494 LITANKEV 501

RESULT 11
Q8Y8K4 PRELIMINARY; PRT; 541 AA.
AC Q8Y8K4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Lmo0897 protein.
GN OrderedLocName=Lmo0897;
OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bioecker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Darvar A., Dehoux P.,

```

```

RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fehri H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hatt J., Jackson D.,
RA Jones L.-M., Kaerst U., Krell T., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:949-952(2001).
DR EMBL: AL581977; CAC8975.1; -.
DR FIR: A1186; A1186.
DR ListList: LMO0897; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0008271; F:sulfate porter activity; IEA.
DR GO: GO:0008272; P:sulfate transport; IEA.
DR InterPro: IPR002645; SPAS.
DR InterPro: IPR01547; Sulph_transp.
DR InterPro: IPR001902; SulP_transp.
DR Pfam: PF01740; STAS_1.
DR Pfam: PF00916; Sulfate_transp. 1.
DR TIGRfam: TIGR00815; sulP_1.
DR PROSITE: PS50801; STAS_1.
KW Complete proteome.
SQ SEQUENCE 541 AA; 58261 MW; B3A3A1A76A36F2A CRC64;

Query Match
Best Local Similarity 76.2%; Score 32; DB 2; Length 541;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CY 1 SLTANKEV 9
DB 154 SLTANKEV 162

RESULT 12
Q6ZK85 PRELIMINARY; PRT; 684 AA.
AC Q6ZK85;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DR RNA recognition motif (RRM)-containing protein-like.
GN Name=OJ1134 H03.9-1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhacoidae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003983; BAB08825.1; -.
DR InterPro: IPR005054; RRM_rec_mot.
DR Pfam: PF00076; RRM_1; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
SQ SEQUENCE 684 AA; 75299 MW; 0E3CF5DE072FDF8 CRC64;

Query Match
Best Local Similarity 76.2%; Score 32; DB 2; Length 684;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

CY 1 SLTANKEV 9
DB 486 SLTANKEV 494

RESULT 13
BAB08825 PRELIMINARY; PRT; 684 AA.
AC BAB08825;

```

DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE RNA recognition motif (RNM)-containing protein-like.  
 GN OJ134\_H03-9-1.  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriatridae; Oryzaceae; Oryza; Oryza sativa.  
 OC NCBI\_TaxID=39947;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 8, BAC  
 clone:OJ134\_H03";  
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBD databases.  
 DR EMBL, AP003883, BAD08825.1, .  
 SQ SEQUENCE 684 AA; 7539 MW; 0E3C3F5DE072FDF8F CRC64;

Query Match 76.2%; Score 32; DB 2; Length 684;  
 Best Local Similarity 66.7%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 LTANKEV 9  
 DB 486 SLISNIKL 494

RESULT 14  
 ID Q74CF5 PRELIMINARY; PRT; 124 AA.  
 AC Q74CF5;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=GSU1719;  
 OS Geobacter sulfurreducens.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;  
 OC Geobacteraceae; Geobacter.  
 OC NCBI\_TaxID=35554;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCA / ATCC 51573;  
 RX PubMed=14671304; DOI=10.1126/science.1086727;  
 RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,  
 RA Heidelberg J.F., Wu D., Ward N.L., Beanan M.J., Dodson R.J.,  
 RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,  
 RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,  
 RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,  
 RA Weidman J.F., Khouli H.M., Feldlyum T.V., Utterback T.R.,  
 RA Van Aken S.E., Lovley D.R., Fraser C.M.;  
 RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface  
 RT environments.";  
 RL Science 302:1967-1969(2003).  
 DR EMBL, AE017213; AAR35096.1; .  
 DR TIGR; GSU1719; .  
 KW Hypothetical protein.  
 SQ SEQUENCE 124 AA; 13279 MW; 0098A8E5109A24DA CRC64;

Query Match 73.8%; Score 31; DB 2; Length 124;  
 Best Local Similarity 85.7%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LTANKE 8  
 DB 88 LTANKE 94

RESULT 15  
 AAR35096 PRELIMINARY; PRT; 124 AA.  
 ID AAR35096

AC AAR35096;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN GSU1719.  
 OS Geobacter sulfurreducens.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;  
 OC Geobacteraceae; Geobacter.  
 OC NCBI\_TaxID=35554;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCA / ATCC 51573;  
 RX PubMed=14671304;  
 RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,  
 RA Heidelberg J.F., Wu D., Ward N.L., Beanan M.J., Dodson R.J.,  
 RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,  
 RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,  
 RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,  
 RA Weidman J.F., Khouli H.M., Feldlyum T.V., Utterback T.R.,  
 RA Van Aken S.E., Lovley D.R., Fraser C.M.;  
 RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface  
 RT environments.";  
 RL Science 302:1967-1969(2003).  
 DR EMBL, AE017213; AAR35096.1; .  
 DR TIGR; GSU1719; .  
 KW Hypothetical protein.  
 SQ SEQUENCE 124 AA; 13279 MW; 0098A8E5109A24DA CRC64;

Query Match 73.8%; Score 31; DB 2; Length 124;  
 Best Local Similarity 85.7%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LTANKE 8  
 DB 88 LTANKE 94

Search completed: November 30, 2004, 07:17:32  
 Job time : 70.7692 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 69.4615 Seconds  
(without alignments)  
46.480 Million cell updates/sec

Title: US-10-008-377a-7  
Perfect score: 42  
Sequence: 1 SLTANKEV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues  
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq\_23sep04:\*  
1: geneseqp1908:\*  
2: geneseqp1908:\*  
3: geneseqp20008:\*  
4: geneseqp20015:\*  
5: geneseqp20025:\*  
6: geneseqp20035:\*  
7: geneseqp20035:\*  
8: geneseqp20045:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	AA022989	AA022989 Human p62
2	42	100.0	675	ABG20950	ABG20950 Novel hum
3	42	100.0	680	AA873858	AA873858 Human NPM
4	42	100.0	680	ABG95067	ABG95067 Human tra
5	42	100.0	680	ABU04347	ABU04347 Human exp
6	42	100.0	680	ABU04362	ABU04362 Human exp
7	42	100.0	680	ABU04358	ABU04358 Human exp
8	42	100.0	680	ABG74472	ABG74472 NPM/ALK f
9	42	100.0	1620	AA873857	AA873857 Human ALK
10	42	100.0	1620	AA873857	AA873857 Human ALK
11	42	100.0	1620	AA874471	AA874471 Human ins
12	42	100.0	1620	AA874471	AA874471 Human ALK
13	33	78.6	560	AA870013	AA870013 Tick carb
14	32	76.2	510	AA870013	AA870013 Tobacco f
15	32	76.2	541	AB847568	AB847568 Listeria
16	32	76.2	541	ABU32622	ABU32622 Protein e
17	31	73.8	41	AA873858	AA873858 Human pro
18	31	73.8	121	AA845378	AA845378 Human sec
19	31	73.8	136	AA001369	AA001369 Human pol
20	31	73.8	250	ADP13932	ADP13932 Human end
21	31	73.8	257	ADP13932	ADP13932 Human end
22	31	73.8	277	ADP13940	ADP13940 Human end
23	31	73.8	319	ADP13938	ADP13938 Human end
24	31	73.8	334	ABG16486	ABG16486 Novel hum
25	31	73.8	339	ABU25245	ABU25245 Protein e
			378	ADH88725	ADH88725 Enterococ

26	31	73.8	510	AB848924	AB848924 Listeria
27	31	73.8	510	ABU32635	ABU32635 Protein e
28	31	73.8	940	ADP13937	ADP13937 Human end
29	31	73.8	1250	ADP13935	ADP13935 Human end
30	31	73.8	1335	ADP13936	ADP13936 Human end
31	31	73.8	1683	ADP44519	ADP44519 Mouse kin
32	31	73.8	2247	ADP13933	ADP13933 Human end
33	31	73.8	2383	ADP13934	ADP13934 Human end
34	31	73.8	2545	AB848906	AB848906 Human NOV
35	30	71.4	122	AA845377	AA845377 Gene 38 h
36	30	71.4	145	AB848905	AB848905 Listeria
37	30	71.4	179	AA848905	AA848905 S. epider
38	30	71.4	179	ABU43169	ABU43169 Protein e
39	30	71.4	180	ABP39370	ABP39370 Staphyloc
40	30	71.4	190	AB840628	AB840628 Human ORF
41	30	71.4	190	ABP00040	ABP00040 Human ORF
42	30	71.4	232	ABU02158	ABU02158 S. pneumo
43	30	71.4	232	ADK46663	ADK46663 Streptoco
44	30	71.4	232	ADK47957	ADK47957 Streptoco
45	30	71.4	233	ABU01769	ABU01769 S. pneumo

## ALIGNMENTS

RESULT 1	AA022989	standard; peptide: 9 AA.
ID	AA022989	standard; peptide: 9 AA.
XX	AA022989;	
XX	17-SEP-2003	(first entry)
DE	Human p62-629 ALK-derived HLA-A*0201 restricted CTL epitope peptide.	
XX	HLA-A*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;	
XX	cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;	
XX	oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;	
XX	t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;	
XX	p62-629.	
OS	Homo sapiens.	
XX	WO2003042243-A2.	
XX	22-MAY-2003.	
XX	14-NOV-2002; 2002WO-EP012764.	
XX	15-NOV-2001; 2001US-00008377.	
XX	(NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.	
XX	(STUD-) ST JUDE CHILDREN'S RES HOSPITAL.	
XX	Gambacorti-Passerini C, Passoni L;	
XX	WPI; 2003-441791/41.	
XX	New HLA-A*0201-Binding Anaplastic Lymphoma Kinase (ALK) peptide,	
XX	useful for preparing a composition for treating ALK-positive lymphoma,	
XX	neuroblastoma or ALK-expressing neoplasia.	
XX	Claim 1; Page 7; 33pp; English.	
XX	The invention relates to a novel HLA-A*0201-binding anaplastic lymphoma	
XX	kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a	
XX	cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase	
XX	which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase	
XX	fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell	
XX	lines. More than 50% of ALCL cases possess a t(2;5) chromosomal	
XX	translocation that leads to the expression of the NPM/ALK fusion protein	
XX	which forms a potent oncogene when constitutively activated. Translocated	
XX	ALK is a widely expressed tumour-associated antigen characteristic of ALK	

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytosolic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumor cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilized during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p621-629  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 CC

SO Sequence 9 AA;

Query Match 100.0%; Score 42; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 |||||  
 1 SLTANKEV 9

RESULT 2  
 ABG20950  
 ID ABG20950 standard; protein; 675 AA.

AC ABG20950;  
 DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #20941.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

PA (HYSB-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-633362/73.

DR N-PSDB; AAS85137.

PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

PS Claim 20; SEQ ID NO 51309; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (II) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC

SO Sequence 675 AA;

Query Match 100.0%; Score 42; DB 4; Length 675;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 |||||  
 616 SLTANKEV 624

RESULT 3  
 AAB73858  
 ID AAB73858 standard; protein; 680 AA.

AC AAB73858;

DT 15-MAY-2001 (first entry)

DE Human NPM/ALK fusion protein.

KW Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;  
 KW chromosomal translocation; cancer; NPM/ALK; fusion.

OS Homo sapiens.

PN US6174674-B1.

PD 16-JAN-2001.

PF 19-JUN-1998; 98US-00100089.

PR 03-DEC-1993; 93US-00160861.

PR 12-OCT-1995; 95US-00542363.

PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Morris SW, Look AT;

DR WPI; 2001-243208/25.

DR N-PSDB; AAB76868.

PT Detection of chromosomal rearrangement or translocations present in t(2;  
 PT 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 PT techniques.

PS Claim 1; Fig 2A; 87pp; English.

XX The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 CC

SO Sequence 680 AA;  
 Query Match 100.0%; Score 42; DB 4; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.6;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLTANKEV 9  
 |||||  
 Db 621 SLTANKEV 629

RESULT 4  
 ABG95067  
 ID ABG95067 standard; protein; 680 AA.  
 XX  
 AC ABG95067;  
 XX  
 DT 04-DEC-2002 (first entry)  
 XX  
 DE Human translocation (2; 5)(p23; q35) protein.  
 XX  
 DE Chromosome aberration; oncogenic fusion protein; cancer;  
 KM proliferative disease; cellular protein isoform; heat shock protein 90;  
 KM HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;  
 KM T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;  
 KM acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;  
 KM acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;  
 KM papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;  
 KM rhabdomyosarcoma; synovial sarcoma; viral infection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200269900-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 01-MAR-2002; 2002MO-US006518.  
 XX  
 PR 01-MAR-2001; 2001US-0272751P.  
 XX  
 PA (CONF-) CONFORMA THERAPEUTICS CORP.  
 XX  
 FI Fritz LC, Burrows FJ;  
 XX  
 DR WPI: 2002-698710/75.  
 DR N-PSDB: ABS73246.  
 XX  
 PT Treating genetically-defined disease associated with chromosomal  
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative  
 PT diseases, involves administering an inhibitor of heat shock protein 90.  
 XX  
 PS Disclosure; page 200-202; 389pp; English.  
 XX  
 CC The invention describes a method of treating genetically-defined disease  
 CC associated with chromosomal aberrations yielding oncogenic fusion  
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous  
 CC cell population, treating proliferative diseases associated with mutant  
 CC protein or cellular protein isoforms (II) dependent on heat shock protein  
 CC (HSP)-90, or selectively treating cells expressing (II) involving  
 CC administering HSP90-inhibitor. The method is useful for treating  
 CC genetically-defined disease with chromosomal aberration yielding  
 CC oncogenic fusion protein, treating cancerous cells containing fusion  
 CC protein in heterogeneous cell population, treating proliferative disease  
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or  
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.  
 CC HSP90), or selectively treating cells expressing mutant protein or cellular  
 CC protein isoform in a patient heterozygous for (II). The method is useful  
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell  
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML,  
 CC or a disease characterised by a solid tumour such as papillary thyroid  
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and  
 CC synovial sarcoma. The method is also useful for treating viral  
 CC infections. This represents a protein encoded by the DNA sequence of a  
 CC chromosome aberration  
 XX  
 SQ Sequence 680 AA;

Query Match 100.0%; Score 42; DB 5; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLTANKEV 9  
 |||||  
 Db 621 SLTANKEV 629

RESULT 5  
 ABU04347  
 ID ABU04347 standard; protein; 680 AA.  
 XX  
 AC ABU04347;  
 XX  
 DT 29-JAN-2003 (first entry)  
 XX  
 DE Human expressed protein tag (EPT) #1013.  
 XX  
 DE Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KM protease; protease inhibitor; transporter; cytoskeletal protein;  
 KM receptor; transcription factor; cancer; MHC;  
 KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200278524-A2.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 28-MAR-2002; 2002MO-US009671.  
 XX  
 PR 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 PA (ZYCO-) ZYCO INC.  
 XX  
 PI Chiciz RM, Tomlinson AJ, Urban RG;  
 XX  
 DR WPI: 2003-040607/03.  
 XX  
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukaemia.  
 XX  
 PS Example 2, SEQ ID NO 1013; 134pp; English.  
 XX  
 CC The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polypeptides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIP0 at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 680 AA;

Query Match 100.0%; Score 42; DB 6; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 621 SLTANKEV 629

RESULT 6  
ID ABU04362 standard; protein; 680 AA.  
XX  
AC ABU04362;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1028.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WC020278524-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 28-MAR-2002; 2002MO-US009671.  
XX  
PR 28-MAR-2001; 2001US-0279495P.  
XX  
PR 21-MAY-2001; 2001US-0292544P.  
XX  
PR 08-AUG-2001; 2001US-0310801P.  
XX  
PR 01-OCT-2001; 2001US-0326370P.  
XX  
PR 04-DEC-2001; 2001US-0336780P.  
XX  
PR 20-FEB-2002; 2002US-0358985P.  
XX  
PA (ZYCO-) ZYCOS INC.  
XX  
PI Chicx RM, Tomlinson AJ, Urban RG;  
XX  
DR WPI; 2003-040607/03.  
XX  
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukaemia.  
XX  
PS Example 2; SEQ ID NO 1028; 134P; English.  
XX  
XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this  
XX polypeptide, is useful for treating cancer. The polypeptide is also  
XX useful for identifying compounds that binds to a naturally processed  
XX class I or class II MHC-binding polypeptide. The polypeptides and  
XX polynucleotides are particularly useful for treating or preventing  
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
XX lymphoma or leukaemia. These are also useful for screening agents for  
XX treating the above mentioned diseases. This sequence represents an  
XX expressed protein tag (EPT) isolated from human tissue for translational  
XX profiling. Note: This sequence does not appear in the printed  
XX specification but was obtained in electronic format directly from WIP0 at  
XX ftp.wipo.int/pub/published\_pcl\_sequences  
SQ Sequence 680 AA;

Query Match 100.0%; Score 42; DB 6; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 621 SLTANKEV 629

RESULT 7  
ID ABU04358 standard; protein; 680 AA.  
XX  
AC ABU04358;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1024.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WC020278524-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 28-MAR-2002; 2002MO-US009671.  
XX  
PR 28-MAR-2001; 2001US-0279495P.  
XX  
PR 21-MAY-2001; 2001US-0292544P.  
XX  
PR 08-AUG-2001; 2001US-0310801P.  
XX  
PR 01-OCT-2001; 2001US-0326370P.  
XX  
PR 04-DEC-2001; 2001US-0336780P.  
XX  
PR 20-FEB-2002; 2002US-0358985P.  
XX  
PA (ZYCO-) ZYCOS INC.  
XX  
PI Chicx RM, Tomlinson AJ, Urban RG;  
XX  
DR WPI; 2003-040607/03.  
XX  
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukaemia.  
XX  
PS Example 2; SEQ ID NO 1024; 134P; English.  
XX  
XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this  
XX polypeptide, is useful for treating cancer. The polypeptide is also  
XX useful for identifying compounds that binds to a naturally processed  
XX class I or class II MHC-binding polypeptide. The polypeptides and  
XX polynucleotides are particularly useful for treating or preventing  
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
XX lymphoma or leukaemia. These are also useful for screening agents for  
XX treating the above mentioned diseases. This sequence represents an  
XX expressed protein tag (EPT) isolated from human tissue for translational  
XX profiling. Note: This sequence does not appear in the printed  
XX specification but was obtained in electronic format directly from WIP0 at  
XX ftp.wipo.int/pub/published\_pcl\_sequences  
SQ Sequence 680 AA;

Query Match 100.0%; Score 42; DB 6; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 |||||  
 DB 621 SLTANKEV 629

RESULT 8  
 ABG74472 ID ABG74472 standard; protein; 680 AA.  
 XX AC ABG74472;  
 XX DT 11-APR-2003 (first entry)  
 XX DE NPM/ALK fusion construct SEQ ID 4.  
 XX KW ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;  
 XX KW t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;  
 XX KW nucleolar phosphoprotein; centromeric; telomeric.  
 XX OS Synthetic.

XX FH Key Location/Qualifiers  
 FT Protein 1..117  
 FT /note="NPM-derived protein fragment"  
 FT Domain 104..115  
 FT /note="potential metal binding domain"  
 FT Protein 118..680  
 FT /note="ALK-derived protein fragment"  
 FT Domain 182..437  
 FT /label= ALK\_catalytic\_domain

XX UN 03-DEC-1997-BI.  
 XX PD 17-SEP-2002.  
 XX PF 28-SEP-2000; 2000US-00670827.  
 XX PR 03-DEC-1993; 93US-00160861.  
 XX PR 12-OCT-1995; 95US-00542363.  
 XX PR 19-JUN-1998; 98US-00100089.

XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX Morris SW, Look AT;  
 XX WPI: 2003-101739/09.  
 XX DR N-PSDB; ABQ77107.  
 XX PT Kit for use in method of detecting t(2;5) chromosomal rearrangements or  
 PT rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic  
 PT lymphoma kinase (ALK) comprises probes that hybridize with NPM and ALK  
 PT genes.

XX Example 1; Fig 2A; 87pp; English.  
 XX This invention describes a novel kit for use in a method of detecting t(2;  
 CC /5) chromosomal rearrangements (CR) or CRs involving nucleolar  
 CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method  
 CC comprises labelled probes that hybridize to a sequence of 2442 or 6226  
 CC nucleotides or their complement, and to region of human chromosome 5/2  
 CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to  
 CC nucleotide). This sequence represents a polypeptide sequence described in  
 CC the disclosure of the invention

XX SQ Sequence 680 AA;

Query Match 100.0%; Score 42; DB 6; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 |||||  
 DB 621 SLTANKEV 629

RESULT 9  
 AAB73857 ID AAB73857 standard; protein; 1620 AA.  
 XX AC AAB73857;  
 XX DT 15-MAY-2001 (first entry)  
 XX DE Human ALK protein.  
 XX KW Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;  
 XX KW chromosomal translocation; cancer.

XX OS Homo sapiens.  
 XX UN 03-DEC-1993; 93US-00160861.  
 XX PR 12-OCT-1995; 95US-00542363.  
 XX PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX PI Morris SW, Look AT;  
 XX WPI: 2001-243208/25.  
 XX DR N-PSDB; AAF76867.

XX PT Detection of chromosomal rearrangement or translocations present in t(2;  
 PT 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 PT techniques.  
 XX Claim 10; Fig 3B; 87pp; English.

XX The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer

XX SQ Sequence 1620 AA;

Query Match 100.0%; Score 42; DB 4; Length 1620;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 |||||  
 DB 1561 SLTANKEV 1569

RESULT 10  
 AA018498 ID AA018498 standard; protein; 1620 AA.  
 XX AC AA018498;

```

DE      11-OCT-2002 (first entry)
XX      Human insulin receptor signaling modifier SEQ ID NO: 18.
XX      Human, insulin receptor signaling; insulin receptor signaling modifier;
XX      ISM; diabetes; metabolic syndrome; antidiabetic.
OS      Homo sapiens.
XX      WO200255664-A2.
XX      18-JUL-2002.
XX      11-JAN-2002; 2002WO-US001048.
XX      12-JAN-2001; 2001US-0261262P.
XX      12-JAN-2001; 2001US-0261303P.
XX      12-JAN-2001; 2001US-0261304P.
XX      12-JAN-2001; 2001US-0261305P.
XX      12-JAN-2001; 2001US-0261336P.
XX      12-JAN-2001; 2001US-0261336P.
XX      12-JAN-2001; 2001US-0261361P.
XX      12-JAN-2001; 2001US-0261456P.
XX      12-JAN-2001; 2001US-0261457P.
XX      12-JAN-2001; 2001US-0261458P.
XX      12-JAN-2001; 2001US-0261459P.
XX      12-JAN-2001; 2001US-0261461P.
XX      12-JAN-2001; 2001US-0261516P.
XX      12-JAN-2001; 2001US-0261531P.
XX      12-JAN-2001; 2001US-0261532P.
XX      12-JAN-2001; 2001US-0261589P.
XX      12-JAN-2001; 2001US-0261590P.
XX      12-JAN-2001; 2001US-0261694P.
XX      12-JAN-2001; 2001US-0261695P.
XX      12-JAN-2001; 2001US-0261697P.
XX      (EXEL-) EXELIXIS INC.
XX      Seidel-Dugan C, Ferguson KC, Kidd T;
XX      WPI; 2002-599664/64.
XX      N-PDB; AAL48617.
XX      Identifying an insulin receptor signaling modulator, useful as drug
XX      targets for treating diabetes or metabolic disorders, comprises
XX      contacting an assay system comprising insulin receptor signaling
XX      modifiers with a test agent.
XX      Disclosure; Page 59-66; 232p; English.
XX      The present invention relates to a method of identifying a candidate
XX      insulin receptor (INR) signaling modulating agent, involving contacting
XX      an assay system comprising an insulin receptor signaling modifier (ISM)
XX      polypeptide or nucleic acid with a test agent, and detecting a test agent
XX      -biased activity of the assay system. The method is useful for
XX      identifying candidate INR signaling modulating agents. ISM genes may be
XX      used as drug targets for treatment of disorders related to INR signaling
XX      polypeptides are useful for identifying and testing agents that modulate
XX      ISM function and for other applications related to the involvement of ISM
XX      in INR signaling, and for identifying subjects having a predisposition to
XX      CC diseases associated with INR signaling. The present sequence is an
XX      ISM protein described in the exemplification of the invention
XX      SQ      Sequence 1620 AA;
XX      Query Match      100.0%; Score 42; DB 5; Length 1620;
XX      Best Local Similarity 100.0%; Pred. No. 4.e;
XX      Matches      9; Conservative      0; Mismatches      0; Gaps      0,
XX      1 SLPANKKEV 9
XX      |||||
XX      1561 SLPANKKEV 1569

```

```

RESULT 11
ID ABG74471 standard; protein, 1620 AA.
XX
XX ABG74471;
AC
XX
XX
XX 11-APR-2003 (first entry)
XX Human ALK protein SEQ ID 2.
DE
XX
XX ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;
XX t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;
XX nucleolar phosphoprotein; centromeric; telomeric.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX Peptide 1..26
XX FT /label=signal_peptide
XX FT 27..1620
XX FT /label=mature_ALK
XX FT 1031..1058
XX Domain
XX FT /label=transmembrane_domain
XX FT 1058..1059
XX FT /note="NPM-ALK fusion junction"
XX FT 1123..1377
XX Domain /label=tyrosine_kinase_catalytic_domain
XX
XX US6451997-B1.
XX
XX 17-SEP-2002.
XX
XX 26-SEP-2000; 2000US-00670827.
XX
XX 03-DEC-1993; 93US-00160861.
XX 12-OCT-1995; 95US-00542363.
XX 19-JUN-1998; 98US-00100089.
XX
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Morris SW, Lock AT;
XX
XX WPI: 2003-101739/09.
XX N-PSDB; ABQ77106.
XX
XX Kit for use in method of detecting t(2;5) chromosomal rearrangements or
XX rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic
XX lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK
XX genes.
XX
XX Example 1; Fig 3B; 87pp; English.
XX
XX This invention describes a novel Kit for use in a method of detecting t(2
XX ;5) chromosomal rearrangements (CR) or CRs involving nucleolar
XX phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method
XX comprises labelled probes that hybridize to a sequence of 2442 or 6226
XX nucleotides or their complement, and to region of human chromosome 5/2
XX having wild-type NPM/ALK genes (the region is centromeric/telomeric to
XX nucleotide). This sequence represents a polypeptide sequence described in
XX the disclosure of the invention
XX
XX Sequence 1620 AA;
SQ
Query Match 100.0%; Score 42; DB 6; Length 1620;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
1 SLTANKREV 9
|||||
1561 SLTANKREV 1569
DB

```

CC	AA770013 standard; protein; 660 AA.
AC	AA770013;
XX	
XX	
DT	25-MAR-2003 (revised)
DT	25-SEP-1995 (first entry)
XX	
DE	Tick carboxypeptidase.
XX	
KV	
XX	
OS	Tick; antigen; carboxypeptidase; vaccine.
PH	Boophilus microplus.
FT	Key
FT	Location/Qualifiers
FT	1..29
FT	/label= signal
FT	30..55
FT	/label= BM91 peptide
FT	185..201
FT	/label= peptide AAT91251
FT	202..209
FT	/label= AAT91141
FT	225..253
FT	/label= T9118
FT	262..269
FT	/label= Bm91 peptide
FT	442..452
FT	/label= T9123(a)
FT	456..481
FT	/label= T9123(b)
FT	599..606
FT	/label= T9109
FT	639..655
FT	/label= C-terminal transmembrane domain
FT	/note= "putative"
XX	
PN	W09504827-A1.
PD	16-FEB-1995.
XX	
PF	10-AUG-1994; 94WO-AU000463.
PR	10-AUG-1993; 93AU-00000458.
XX	
PA	(BIOT-) BIOTECH AUSTRALIA PTY LTD.
PA	(CSIR ) COMMONWEALTH SCI & IND RES ORG.
PI	Cobon GS, Willadsen P, Kemp DH, Tellam RL;
DR	WP1; 1995-090905/12.
XX	N-PsDBj; AAQ82948.
PT	New antigenic tick carboxypeptidase and corresp. DNA - are used in vaccines for producing antibodies against ticks, insects and nematodes.
PS	Disclosure; Fig 6; 139pp; English.
CC	Clone A5 was prepd. from adult tick cDNA library. Clone 4U1 was prepd. from the larval stage of B. microplus (Calliope strain). AAQ82948 is a hybrid of sequences from clone 4U1 (nt 1-966 & 1747-2047) and A5 (nts 967-1746). The translation of the tick carboxypeptidase cDNA sequence is shown in AA770013. All the native tick carboxypeptidase sequences listed in Table 11 (see AA770014-R70023) are found in the translation (see FT).
CC	The predicted AA sequence agrees with the peptide sequence for all peptides with 2 exceptions. These differences are Asp for Glut14 in peptide T9126, and Asn for Asp12 in peptide T918. Tick carboxypeptidase has a mol. wt. 75172. In addition to the features in FT, the carboxypeptidase AA sequence has 8 potential N-linked glycosylation sites and a potential glycosylphosphatidylinositol anchor sequence similar to that found in Embe6. It has significant homology with zinc dependent dipeptidyl carboxypeptidases from mammals. (Updated on 25-MAR-2003 to

```

CC correct FN field.)
XX
SQ Sequence 660 AA;
Query Match 78.6%; Score 33; DB 2; Length 660;
Best Local Similarity 87.5%;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps
QY 2 LTANKEV 9
|||
173 LTRNKEV 180
DB
RESULT 13
AA05901
ID AA05901 standard; protein; 510 AA.
AC
XX AA05901;
XX
XX 02-AUG-1999 (first entry)
XX
DE Tobacco fatty acid hydroxylase CYP94A6.
XX
XX Fatty acid omega-hydroxylase; Cytochrome P450; transgenic plant; lipid;
KM hydroxylation; epoxidation; oilseed; vegetable oil; crop protection;
KW omega-hydroxy acid; CYP94A6; tobacco.
XX
XX Nicotiana tabacum.
XX
XX Key Location/Qualifiers
FH 317..328
FT Domain /note="haem-binding domain, corresponds to signature
FT motif claimed in Claim 2"
FN
XX W09918224-A1.
XX
XX 15-APR-1999.
XX
XX 06-OCT-1998; 98WO-IB001716.
XX
XX 06-OCT-1997; 97US-0060960P.
XX
XX (CNRS ) CENT NAT RECH SCI.
XX
XX Tillet N, Pinot F, Benveniste I, Le Bouguin R, Helvig C, Batard Y,
PI Cabello-Ruattado F, Werck-Reichhart D, Salauin J, Durst F,
XX WPI; 1999-264030/22.
XX
XX N-PSDB; AAX58405.
XX
XX Nucleic acid encoding plant fatty acid hydroxylases.
XX
XX Example 7, Fig 26A-B; 157bp; English.
XX
XX This sequence represents a cytochrome P450 protein, CYP94A6, of tobacco
CC that is expected to have fatty acid hydroxylase activity since it
CC displays the characteristic signature motif for this class of enzymes.
CC The invention provides isolated nucleic acids (see AAX58400-06) encoding
CC plant fatty acid hydroxylases (see AAX05896-902). Also claimed are host
CC cells, transgenic plants and compositions consisting of the plant fatty
CC acid hydroxylase, a process for isolating additional fatty acid
CC hydroxylase genes from a plant, and a process of altering fatty acid
CC composition in a plant by expressing the plant fatty acid hydroxylase in
CC a transgenic plant, and hydroxylating or epoxidating a fatty acid
CC substrate in the plant. Manipulating the hydroxylated fatty acid content
CC of plants will modify resistance to drought and attack by insects and
CC other pests. The transgenic plants may also be used as sources of
CC hydroxylated and epoxidized fatty acids useful in the manufacture of e.g
CC lubricants, anti-slip agents, plasticisers, coating agents, detergents
CC and surfactants
XX
XX Sequence 510 AA.
SQ

```

Query Match 76.2%; Score 32; DB 2; Length 510;  
 Best Local Similarity 87.5%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKEV 9  
 DB 494 LTANKEV 501

RESULT 14  
 ABBA7568  
 ID ABBA7568 standard; protein; 541 AA.

AC ABBA7568;

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #272.

KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
 KM vitamin B12; bacterial infection; disease.

OS Listeria monocytogenes.

PN WO200177335-A2.

PD 18-OCT-2001.

PF 11-APR-2001; 2001WO-FR001118.

PR 11-APR-2000; 2000FR-00004629.

XX (INSP) INST PASTEUR.

XX Buchrieser C, Frangeul L, Couve E, Rusnick C, Felhi H, Deboux P,  
 PI Dusunreyet O, Cretotant F, Nedjari H, Glaser P, Kunst F, Cossart P,  
 PI Dangles J, Godel W, Xieft J, Kuhn M, Ng E, Vazquez-Boland JA,  
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A,  
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L,  
 PI Perez-Diaz J, Baguerio F, Garcia Del Portillo F, Gomez-Lopez N,  
 PI Madueno E, De Pablos B, Wehlend J, Kaerst U, Entian K, Hauf J,  
 PI Rose M, Voss H;  
 XX WPI: 2002-010914/01.

PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment  
 PT and prevention of Listeria and related bacterial infections, and related  
 PT polypeptides.

PS Claim 6; SEQ ID NO 273; 192pp; French.

CC The present invention relates to the genome sequence of Listeria  
 CC monocytogenes EGD-e (see ABBA03041). The genome sequence and fragments of  
 CC it are useful for selecting probes and primers for detecting genes in L.  
 CC monocytogenes and related organisms, and for studying genetic  
 CC polymorphisms and other genomes. The present sequence is a protein  
 CC encoded by the genome sequence of the present invention. Proteins  
 CC expressed from the genome sequence are useful for raising specific  
 CC antibodies. Identification of L. monocytogenes and related organisms, and  
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin  
 CC B12. The genome sequence and proteins encoded by it are also useful for  
 CC selecting compounds that regulate gene expression and cell replication  
 CC and modulate L. monocytogenes-related diseases. In addition, the genome  
 CC sequence and proteins encoded by it are useful in pharmaceutical and  
 CC vaccine compositions for the treatment or prevention of infections by L.  
 CC monocytogenes and related organisms. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pat\_sequences

XX Sequence 541 AA;

Query Match 76.2%; Score 32; DB 5; Length 541;

Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SFTANKEV 9  
 DB 154 SFTANKEV 162

RESULT 15  
 ABU32622  
 ID ABU32622 standard; protein; 541 AA.

AC ABU32622;

DT 19-JUN-2003 (first entry)

DE Protein encoded by prokaryotic essential gene #1819.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Listeria monocytogenes.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342933P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362639P.

XX (ELIT) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen XL, Zyskind JW,  
 PI Wall D, Traxler JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI: 2003-023926/02.

PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids, required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 60546; 176pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway of  
 CC the gene product or that inhibits cellular proliferation; (8)  
 CC required for proliferation, or that inhibits cellular proliferation; (9)  
 CC identifying a gene required for cellular proliferation or its gene product lies  
 CC pathway in which a proliferation-regulated gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids



CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 541 AA:

Query Match 76.2%; Score 32; DB 6; Length 541;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
Db 154 SPVANNKEI 162

Search completed: November 30, 2004, 07:39:59  
Job time : 71.715 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 30, 2004, 07:19:22 ; Search time 339.231 Seconds  
(without alignments)  
9.410 Million cell updates/sec

Title: US-10-008-377A-7

Perfect score: 42

Sequence: 1 SLTANKEV 9

## Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	42	100.0	9 US-10-008-377-7	Sequence 7, Appl1
2	42	100.0	680 US-09-827-949-4	Sequence 4, Appl1
3	42	100.0	1620 US-09-827-949-2	Sequence 2, Appl1
4	35	83.3	268 US-10-424-599-202343	Sequence 202343,
5	33	78.6	96 US-10-425-115-343624	Sequence 343624,
6	33	78.6	1239 US-10-437-963-149396	Sequence 149396,
7	32	76.2	210 US-10-437-963-195871	Sequence 195871,
8	32	76.2	541 US-10-282-122A-60546	Sequence 60546, A
9	31	73.8	41 US-10-282-115-873	Sequence 873, App
10	31	73.8	116 US-10-425-114-43189	Sequence 43189, A
11	31	73.8	162 US-10-425-115-242109	Sequence 242109,
12	31	73.8	202 US-10-424-599-211762	Sequence 211762,
13	31	73.8	339 US-10-282-122A-53169	Sequence 53169, A

14	31	73.8	367 US-10-424-599-178778	Sequence 178778,
15	31	73.8	510 US-10-282-122A-60559	Sequence 60559, A
16	31	73.8	1594 US-10-425-115-291526	Sequence 291526,
17	31	73.8	2545 US-10-042-865-12	Sequence 12,
18	30	71.4	179 US-10-283-71093	Sequence 71093, A
19	30	71.4	285 US-10-365-493-16391	Sequence 16391, A
20	30	71.4	297 US-10-425-115-425348	Sequence 425348,
21	30	71.4	543 US-09-871-212-7	Sequence 7, Appl1
22	30	71.4	608 US-10-282-122A-71118	Sequence 71118, A
23	30	71.4	608 US-10-381-779-30	Sequence 30, Appl
24	30	71.4	608 US-10-369-493-1058	Sequence 3058, Ap
25	30	71.4	868 US-10-437-963-123248	Sequence 123248,
26	30	71.4	1082 US-10-437-963-167162	Sequence 167162,
27	30	71.4	1734 US-09-852-027-81	Sequence 81, Appl
28	30	71.4	1734 US-10-377-035-74	Sequence 41, Appl
29	30	71.4	1734 US-10-042-865-52	Sequence 82, Appl
30	30	71.4	1798 US-10-311-034-23	Sequence 23, Appl
31	30	71.4	1798 US-10-618-941-70	Sequence 70, Appl
32	30	71.4	1805 US-10-425-114-54369	Sequence 54369, A
33	29	69.0	68 US-10-425-115-232919	Sequence 232919,
34	29	69.0	147 US-10-767-701-58957	Sequence 58957, A
35	29	69.0	154 US-10-428-115-293462	Sequence 293462,
36	29	69.0	159 US-10-767-701-62395	Sequence 62395, A
37	29	69.0	178 US-10-424-599-167198	Sequence 167198,
38	29	69.0	202 US-10-767-701-42169	Sequence 42169, A
39	29	69.0	222 US-10-312-273-339	Sequence 339, App
40	29	69.0	222 US-10-289-762-783	Sequence 783, App
41	29	69.0	283 US-10-424-599-776904	Sequence 776904,
42	29	69.0	334 US-10-176-306-38	Sequence 38, Appl
43	29	69.0	342 US-10-767-701-44976	Sequence 44976, A
44	29	69.0	389 US-10-282-122A-74587	Sequence 74587, A
45	29	69.0	482 US-10-365-493-1819	Sequence 1819, Ap

## ALIGNMENTS

RESULT 1  
US-10-008-377-7  
; Sequence 7, Application US/10006377  
; Publication No. US20030157101A1  
; GENERAL INFORMATION:  
; APPLICANT: Gambacorti-Passerini, Carlo  
; APPLICANT: Passoni, Lorena  
; TITLE OF INVENTION: Immunogenic Alk Peptides  
; FILE REFERENCE: 045922/241203  
; CURRENT APPLICATION NUMBER: US/10/008,377  
; CURRENT FILING DATE: 2001-11-15  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq For Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-008-377-7

Query Match 100.0%; Score 42; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLTANKEV 9  
Db 1 SLTANKEV 9  
RESULT 2  
US-09-827-949-4  
; Sequence 4, Application US/09827949  
; Patent No. US20010021505A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Stephan W.  
; APPLICANT: Look, A. Thomas  
; TITLE OF INVENTION: Alk Protein Tyrosine Kinase/Receptor and Ligands Thereof

FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/570,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 4  
LENGTH: 680  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-4

Query Match 100.0%; Score 42; DB 9; Length 680;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 621 SLTANKEV 629

RESULT 3  
US-09-827-949-2  
Sequence 2, Application US/09827949  
Patent No. US20010021505A1  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/570,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 2  
LENGTH: 1620  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-2

Query Match 100.0%; Score 42; DB 9; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 1561 SLTANKEV 1569

RESULT 4  
US-10-424-599-202343  
Sequence 202343, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 202343  
LENGTH: 268  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_24740C.1.pep  
US-10-424-599-202343

Query Match 83.3%; Score 35; DB 15; Length 268;  
Best Local Similarity 77.8%; Pred. No. 39;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 175 SLTANKEV 183

RESULT 5  
US-10-425-115-343624  
Sequence 343624, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 343624  
LENGTH: 96  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_76547C.1.pep  
US-10-425-115-343624

Query Match 78.6%; Score 33; DB 17; Length 96;  
Best Local Similarity 66.7%; Pred. No. 33;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 13 SLTANKEV 21

RESULT 6  
US-10-437-963-149396  
Sequence 149396, Application US/10437963  
Publication No. US20040123943A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbaruk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14

```
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149396
; LENGTH: 1239
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) - (1239)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_49733C.1.Pep
; US-10-437-963-149396

Query Match
Best Local Similarity 78.6%; Score 33; DB 16; Length 1239;
Pred. No. 5.8e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKE 8
Db 519 SLTANKE 526

RESULT 7
; Sequence 195871, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barabuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195871
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91778C.1.Pep
; US-10-437-963-195871

Query Match
Best Local Similarity 76.2%; Score 32; DB 16; Length 210;
Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
Db 7 SLTANKEV 15

RESULT 8
; Sequence 60546, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malore, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Olsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
```

```
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60546
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
; US-10-282-122A-60546

Query Match
Best Local Similarity 76.2%; Score 32; DB 15; Length 541;
Pred. No. 3.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
Db 154 SLTANKEV 162

RESULT 9
; Sequence 873, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hysed Inc
; TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 873
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-296-115-873

Query Match
Best Local Similarity 73.8%; Score 31; DB 15; Length 41;
Pred. No. 33;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
Db 12 SLTANKEV 20
```

## RESULT 10

US-10-425-114-43189  
; Sequence 43189, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 43189  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700748451\_FLI.pep  
US-10-425-114-43189

Query Match 73.8%; Score 31; DB 15; Length 116;  
Best Local Similarity 87.5%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANMKE 8  
|||  
DB 77 SLTANMKE 84

## RESULT 11

US-10-425-115-242109  
; Sequence 242109, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 242109  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(162)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MWT4577\_152386C.1.pep  
US-10-425-115-242109

Query Match 73.8%; Score 31; DB 17; Length 162;  
Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANMKEV 9  
|||  
DB 116 SLTANMKEV 124

## RESULT 12

US-10-424-559-211762

Sequence 211762, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 211762  
; LENGTH: 202  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(202)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MWT3847\_33248C.1.pep  
US-10-424-559-211762

Query Match 73.8%; Score 31; DB 15; Length 202;  
Best Local Similarity 87.5%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANMKE 8  
|||  
DB 163 SLTANMKE 170

## RESULT 13

US-10-282-122A-53169  
; Sequence 53169, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trewick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 53169  
 LENGTH: 339  
 TYPE: PRT  
 ORGANISM: Clostridium difficile  
 US-10-282-122A-53169

Query Match 73.8%; Score 31; DB 15; Length 339;  
 Best Local Similarity 66.7%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 ||:||||  
 Db 98 SLTANKEV 106

RESULT 14  
 US-10-424-599-178778  
 Sequence 178778, Application US/10424599  
 Publication No. US20040031072A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J  
 APPLICANT: Kovacic David K  
 APPLICANT: Zhou Yihua  
 APPLICANT: Cao Yongwei  
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53223) B  
 CURRENT APPLICATION NUMBER: US/10/424,599  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 285664  
 SEQ ID NO 178778  
 LENGTH: 367  
 TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (1) - (367)  
 OTHER INFORMATION: unsure at all Xaa locations  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_WRT3847\_132453C.1.Pep  
 US-10-424-599-178778

Query Match 73.8%; Score 31; DB 15; Length 367;  
 Best Local Similarity 75.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKE 8  
 ||:||||  
 Db 305 SLTANKE 312

RESULT 15  
 US-10-282-122A-60559  
 Sequence 60559, Application US/10282122A  
 Publication No. US20040029129A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haseibeck, Robert  
 APPLICANT: Ohlsen, Kari  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: EPIGRA.034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/151,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269,308  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 60559  
 LENGTH: 510  
 TYPE: PRT  
 ORGANISM: Listeria monocytogenes  
 US-10-282-122A-60559

Query Match 73.8%; Score 31; DB 15; Length 510;  
 Best Local Similarity 55.6%; Pred. No. 5.7e+02;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 ||:||||  
 Db 224 TITANMKDL 232

Search completed: November 30, 2004, 08:06:37  
 Job time : 340.231 secs

**This Page Blank (uspio)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 22.6154 Seconds  
(without alignments)  
26,392 Million cell updates/sec

Title: US-10-008-377A-7  
Sequence: 1 SLTANKREV 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents RA: \*  
1: /cgn2\_6/prodata/1/1aa/5A.CONB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B.CONB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A.CONB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B.CONB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCITUS.CONB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	42	100.0	US-08-542-363-4	Sequence 4, App1
2	42	100.0	US-09-100-089-4	Sequence 4, App1
3	42	100.0	US-09-670-827-4	Sequence 4, App1
4	42	100.0	US-09-827-943-4	Sequence 4, App1
5	42	100.0	US-08-542-363-2	Sequence 2, App1
6	42	100.0	US-09-100-089-2	Sequence 2, App1
7	42	100.0	US-09-670-827-2	Sequence 2, App1
8	42	100.0	US-09-827-943-2	Sequence 2, App1
9	31	73.8	US-09-134-000C-6610	Sequence 6610, Ap
10	30	71.4	US-09-134-001C-1808	Sequence 1808, Ap
11	30	71.4	US-09-134-001C-4215	Sequence 4215, Ap
12	30	71.4	US-09-583-110-3178	Sequence 3178, Ap
13	30	71.4	US-09-583-110-4472	Sequence 4472, Ap
14	30	71.4	US-09-252-991A-16966	Sequence 16966, A
15	30	71.4	US-09-710-279-380	Sequence 320, App
16	30	71.4	US-09-328-352-5957	Sequence 5957, Ap
17	30	71.4	US-09-134-000C-4057	Sequence 4057, Ap
18	30	71.4	US-09-710-279-232	Sequence 232, App
19	30	71.4	US-09-248-796A-14512	Sequence 14512, A
20	30	71.4	US-09-134-001C-4248	Sequence 4248, Ap
21	29	69.0	US-09-198-452A-783	Sequence 783, App
22	29	69.0	US-09-328-352-4801	Sequence 4801, Ap
23	29	69.0	US-09-543-681A-7802	Sequence 7802, Ap
24	29	69.0	US-09-071-035-364	Sequence 364, App
25	29	69.0	US-09-370-838-67	Sequence 67, App1
26	29	69.0	US-09-071-035-362	Sequence 362, App
27	29	69.0	US-09-538-092-944	Sequence 944, App

28	29	69.0	764	4	US-09-854-133-67	Sequence 67, App1
29	29	69.0	776	4	US-09-134-000C-5717	Sequence 5717, Ap
30	28	66.7	100	4	US-09-248-796A-26482	Sequence 26482, A
31	28	66.7	104	3	US-08-946-329A-97	Sequence 97, App1
32	28	66.7	104	4	US-09-248-796A-25322	Sequence 25322, A
33	28	66.7	156	4	US-09-540-236-2381	Sequence 2381, Ap
34	28	66.7	189	4	US-09-270-767-35780	Sequence 35780, A
35	28	66.7	189	4	US-09-270-767-50997	Sequence 50997, A
36	28	66.7	251	4	US-09-328-352-4867	Sequence 4867, Ap
37	28	66.7	393	4	US-09-248-796A-18084	Sequence 18084, A
38	28	66.7	510	4	US-09-134-000C-5774	Sequence 5774, Ap
39	28	66.7	521	4	US-09-107-532A-6431	Sequence 6431, Ap
40	28	66.7	709	1	US-07-814-964-7	Sequence 7, App1
41	28	66.7	709	1	US-08-258-444-7	Sequence 7, App1
42	28	66.7	709	1	US-08-328-809-2	Sequence 2, App1
43	28	66.7	709	3	US-09-015-003-2	Sequence 2, App1
44	28	66.7	709	4	US-08-866-840-2	Sequence 2, App1
45	28	66.7	709	4	US-09-538-092-1287	Sequence 1287, Ap

## ALIGNMENTS

RESULT 1  
US-08-542-363-4  
; Sequence 4, Application US/08542363  
; Patent No. 5770421  
; GENERAL INFORMATION:  
; APPLICANT: MORRIS, Stephan W.  
; APPLICANT: LOCK, A. Thomas  
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
; TITLE OF INVENTION: Ligands thereof  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/542,363  
; FILING DATE: 12-Oct-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fox, Samuel L.  
; REGISTRATION NUMBER: 30,353  
; REFERENCE/DOCKET NUMBER: 0656.040001/SLF/GKT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 680 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-542-363-4

Query Match 100.0%; Score 42; DB 1; Length 680;  
Best Local Similarity 100.0%; Pred. No. 0.44; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLTANKREV 9  
D5 621 SLTANKREV 629

RESULT 2  
US-09-100-089-4

Sequence 4, Application US/09100089  
Patent No. 6174674

GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA

ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400002

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 680 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-09-100-089-4

Query Match  
Best Local Similarity 100.0%; Score 42; DB 3; Length 680;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SLTANKEV 9  
|||||  
621 SLTANKEV 629

RESULT 3  
US-09-670-827-4

Sequence 4, Application US/09670827  
Patent No. 6451997

GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400002

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 680 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-09-100-089-4

Query Match  
Best Local Similarity 100.0%; Score 42; DB 3; Length 680;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SLTANKEV 9  
|||||  
621 SLTANKEV 629

RESULT 4  
US-09-827-949-4

Sequence 4, Application US/09827949  
Patent No. 6686548

GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-Sep-2000

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400003

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 680 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-09-670-827-4

Query Match  
Best Local Similarity 100.0%; Score 42; DB 4; Length 680;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SLTANKEV 9  
|||||  
621 SLTANKEV 629

RESULT 4  
US-09-827-949-4

Sequence 4, Application US/09827949  
Patent No. 6686548

GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400002

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 680 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-09-100-089-4

Query Match  
Best Local Similarity 100.0%; Score 42; DB 4; Length 680;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SLTANKEV 9  
|||||  
621 SLTANKEV 629

RESULT 4  
US-09-827-949-4

Sequence 4, Application US/09827949  
Patent No. 6686548

GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400002

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

Db 621 SLTANKREV 629

RESULT 5  
US-08-542-363-2  
Sequence 2, Application US/06542363  
Patent No. 5770421  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,363  
FILING DATE: 12-OCT-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400001/SLF/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1620 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-542-363-2

Query Match 100.0%; Score 42; DB 1; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 1,2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKREV 9  
Db 1561 SLTANKREV 1569

RESULT 6  
US-09-100-089-2  
Sequence 2, Application US/09100089  
Patent No. 6174674  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1620 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-100-089-2

Query Match 100.0%; Score 42; DB 3; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 1,2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKREV 9  
Db 1561 SLTANKREV 1569

RESULT 7  
US-09-670-827-2  
Sequence 2, Application US/09670827  
Patent No. 645197  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-SEP-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
APPLICATION NUMBER: US 09/100,089  
FILING DATE: 19-JUN-1998  
ATTORNEY/AGENT INFORMATION:

```
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0400003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-670-827-2
Query Match 100.0%; Score 42; DB 4; Length 1620;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 SLTANKEV 9
Db 1561 SLTANKEV 1569
RESULT 8
US-09-827-949-2
Sequence 2, Application US/09827949
Patent No. 6696548
GENERAL INFORMATION:
APPLICANT: Morris, Stephan W.
TITLE OF INVENTION: ALK PROTEIN Tyrosine Kinase/Receptor and Ligands Thereof
FILE REFERENCE: 0656.0400004
CURRENT APPLICATION NUMBER: US/09/827,949
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: US 09/670,827
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 09/100,089
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: US 08/542,363
PRIOR FILING DATE: 1995-10-12
PRIOR APPLICATION NUMBER: US 08/160,861
PRIOR FILING DATE: 1993-12-03
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 1620
TYPE: PRT
ORGANISM: Homo sapiens
US-09-827-949-2
Query Match 100.0%; Score 42; DB 4; Length 1620;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 SLTANKEV 9
Db 1561 SLTANKEV 1569
RESULT 9
US-09-134-000C-6610
Sequence 6610, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
```

```
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6610
LENGTH: 378
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-6610
Query Match 73.8%; Score 31; DB 4; Length 378;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
CY 1 SLTANKEV 9
Db 79 SLTANKEV 87
RESULT 10
US-09-710-279-1808
Sequence 1808, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1808
LENGTH: 179
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1808
Query Match 71.4%; Score 30; DB 4; Length 179;
Best Local Similarity 75.0%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CY 1 SLTANKEV 8
Db 51 SLTANKEV 58
RESULT 11
US-09-134-001C-4215
Sequence 4215, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4215
LENGTH: 180
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4215
Query Match 71.4%; Score 30; DB 3; Length 180;
```

Best Local Similarity 75.0%; Pred. No. 47;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 8  
Db 52 SLTANKEV 59

RESULT 12  
US-09-583-110-3178  
Sequence 3178, Application US/09583110  
Patent No. 6699703  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al.  
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
FILE REFERENCE: PATH00-074  
CURRENT APPLICATION NUMBER: US/09/583,110  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/107,433  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/085,131  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: US 60/051,553  
PRIOR FILING DATE: 1997-07-02  
NUMBER OF SEQ ID NOS: 5322  
SEQ ID NO 3178  
LENGTH: 232  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-583-110-3178

Query Match  
Best Local Similarity 71.4%; Score 30; DB 4; Length 232;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKEV 9  
Db 86 LTANKEV 93

RESULT 13  
US-09-583-110-4472  
Sequence 4472, Application US/09583110  
Patent No. 6699703  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al.  
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
FILE REFERENCE: PATH00-074  
CURRENT APPLICATION NUMBER: US/09/583,110  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/107,433  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/085,131  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: US 60/051,553  
PRIOR FILING DATE: 1997-07-02  
NUMBER OF SEQ ID NOS: 5322  
SEQ ID NO 4472  
LENGTH: 232  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-583-110-4472

Query Match  
Best Local Similarity 71.4%; Score 30; DB 4; Length 232;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKEV 9  
Db 86 LTANKEV 93

RESULT 14  
US-09-252-991A-16966  
Sequence 16966, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfeld et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 16966  
LENGTH: 303  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16966

Query Match  
Best Local Similarity 71.4%; Score 30; DB 4; Length 303;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
Db 64 SLTANKEV 72

RESULT 15  
US-09-710-279-320  
Sequence 320, Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMBERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: P03480US  
CURRENT APPLICATION NUMBER: US/09/710,279  
PRIOR FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 320  
LENGTH: 338  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-320

Query Match  
Best Local Similarity 71.4%; Score 30; DB 4; Length 338;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
Db 329 SIRTANKEV 337

Search completed: November 30, 2004, 07:19:17  
Job time : 23.6154 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 08:27:53 / Search time 50 seconds

(without alignments)  
17.319 Million cell updates/sec

Title: US-10-008-377A-7  
Perfect score: 42  
Sequence: 1 SLTANKEV 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 791

Minimum DB seg length: 0  
Maximum DB seg length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: PIR 79:  
1: PIR1:  
2: PIR2:  
3: PIR3:  
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	45.2	8	1	IFSAME
2	16	38.1	6	2	cytochrome P-450
3	15	35.7	5	2	hypoxanthine phosphoribosyl transferase
4	14	33.3	4	2	endothelial nitric oxide synthase
5	14	33.3	8	2	variant surface glycoprotein
6	13	31.0	8	2	glycine N-acyltransferase
7	13	31.0	8	2	phosphoenolpyruvate carboxykinase
8	13	31.0	9	2	phosphoenolpyruvate carboxykinase
9	12	28.6	4	2	phosphoenolpyruvate carboxykinase
10	12	28.6	5	2	phosphoenolpyruvate carboxykinase
11	12	28.6	7	2	phosphoenolpyruvate carboxykinase
12	12	28.6	7	2	phosphoenolpyruvate carboxykinase
13	12	28.6	7	2	phosphoenolpyruvate carboxykinase
14	12	28.6	7	2	phosphoenolpyruvate carboxykinase
15	12	28.6	7	2	phosphoenolpyruvate carboxykinase
16	11	26.2	4	2	phosphoenolpyruvate carboxykinase
17	11	26.2	5	2	phosphoenolpyruvate carboxykinase
18	11	26.2	6	2	phosphoenolpyruvate carboxykinase
19	11	26.2	7	2	phosphoenolpyruvate carboxykinase
20	11	26.2	7	2	phosphoenolpyruvate carboxykinase
21	11	26.2	7	2	phosphoenolpyruvate carboxykinase
22	11	26.2	7	2	phosphoenolpyruvate carboxykinase
23	11	26.2	7	2	phosphoenolpyruvate carboxykinase
24	11	26.2	7	2	phosphoenolpyruvate carboxykinase
25	11	26.2	8	2	phosphoenolpyruvate carboxykinase
26	11	26.2	8	2	phosphoenolpyruvate carboxykinase
27	11	26.2	8	2	phosphoenolpyruvate carboxykinase
28	11	26.2	8	2	phosphoenolpyruvate carboxykinase
29	11	26.2	8	2	phosphoenolpyruvate carboxykinase

30	11	26.2	8	2	SLP288	acylase - Kluver
31	11	26.2	9	2	PT0670	T-cell receptor be
32	11	26.2	9	2	B30572	T-cell receptor be
33	11	26.2	9	2	PH0902	T-cell receptor be
34	11	26.2	9	2	S78426	52.5K protein - sp
35	11	26.2	9	2	S56004	glucan 1,3-beta-gl
36	11	26.2	9	2	S19523	orf AB protein - S
37	11	26.2	9	2	PC7076	spectrin alpha cha
38	11	26.2	9	2	PC7078	unidentified 48.7K
39	11	26.2	9	2	B24362	chloramphenicol O-
40	11	26.2	9	2	S30494	cat gene leader pe
41	10	23.8	3	3	S13894	histidinol dehydro
42	10	23.8	3	3	PT0578	T-cell receptor be
43	10	23.8	4	2	140870	phospholipase C (B
44	10	23.8	4	2	S17255	ribosomal protein
45	10	23.8	5	2	B37325	pap fibribial regul

#### ALIGNMENTS

##### RESULT 1

IFSAME  
probable msrA leader peptide - Staphylococcus epidermidis  
C:Species: Staphylococcus epidermidis  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: S11157  
R:Rosa, J.I.; Bady, E.A.; Cove, J.R.; Cunliffe, W.J.; Baumberg, S.; Wootton, J.C.  
Mol. Microbiol. 4, 1207-1214, 1990  
A:Title: Inducible erythromycin resistance in staphylococci is encoded by a member of the  
A:Reference number: S11157; MIMD:1041730; PMID:2233255  
A:Accession: S11157  
A:Molecule type: DNA  
A:Residues: 1-8 <ROS>  
A:Cross-references: UNIPROT:P23211; EMBL:X52085; NID:947000; PIDN:CA36303.1; PID:958165  
C:Superfamily: probable msrA leader peptide

##### Query Match

Best Local Similarity 50.0%; Pred. No. 2.8e+05; Length 8;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LTANMK 7  
DB 1 MTASMR 6

##### RESULT 2

149424  
cytochrome P-450  
C:Species: Mus musculus (western wild mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: I49424  
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.F.  
Mamm. Genome 5, 349-355, 1994  
A:Title: Genetic mapping of 40 cDNA clones from the mouse genome by PCR.  
A:Reference number: I48934; MIMD:94319082; PMID:8043949  
A:Accession: I49424  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-6 <RES>  
A:Cross-references: EMBL:U05745; NID:9497084; PIDN:AA60481.1; PID:9642831  
C:Keywords: hydrolase; serine proteinase

##### Query Match

Best Local Similarity 38.1%; Score 16; DB 2; Length 6;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 NMK 7  
DB 2 NMK 4

##### RESULT 3

A37114  
hypoxanthine phosphoribosyltransferase (EC 2.4.2.8) - fluke (Schistosoma mansoni) (fragm  
C/Species: Schistosoma mansoni  
C/Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 23-Jun-1993  
C/Accession: A37114  
R/Yuan, L.; Craig, S.P.; McKerrow, J.H.; Wang, C.C.  
J. Biol. Chem. 265, 13528-13532, 1990  
A/Title: The hypoxanthine-guanine phosphoribosyltransferase of Schistosoma mansoni. Funct  
A/Reference number: A37114; MUID:90337955; PMID:2199439  
A/Accession: A37114  
A/Status: Preliminary  
A/Molecule type: protein  
A/Residues: 1-5 <YDA>  
C/Keywords: glycosyltransferase; pentosyltransferase

Query Match 35.7%; Score 15; DB 2; Length 5;  
Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LITNM 6  
DB 1 MSSNM 5

RESULT 4  
endoglucanase F - Clostridium thermocellum (fragment)  
C/Species: Clostridium thermocellum  
C/Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004  
C/Accession: I40804  
R/Mishra, S.; Beguin, P.; Aubert, J.  
J. Bacteriol. 173, 80-85, 1991  
A/Title: Transcription of clostridium thermocellum endoglucanase genes celf and celd.  
A/Reference number: I40804; MUID:91100322; PMID:1987137  
A/Accession: I40804  
A/Status: Preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-4 <RES>  
A/Cross-references: UNIPROT:P26224; GB:M64363; NID:G144771  
C/Genetics:  
A/Genes: celf  
A/Start codon: TTG

Query Match 33.3%; Score 14; DB 2; Length 4;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 1 MKKI 4

RESULT 5  
A21440  
variant surface glycoprotein pSLc1 - Trypanosoma brucei (fragment)  
C/Species: Trypanosoma brucei  
C/Date: 19-Nov-1988 #sequence\_revision 19-Nov-1988 #text\_change 09-Jul-2004  
C/Accession: A21440  
R/Harsons, W.; Nelson, R.G.; Watkins, K.P.; Agabian, N.  
Cell 38, 309-316, 1984  
A/Title: Trypanosome mRNAs share a common 5' spliced leader sequence.  
A/Reference number: A90853; MUID:84282716; PMID:6088073  
A/Accession: A21440  
A/Molecule type: mRNA  
A/Residues: 1-8 <PAR>  
A/Cross-references: UNIPROT:P22225; GB:K02195; NID:9162150; PID:9162151  
C/Keywords: glycoprotein

Query Match 33.3%; Score 14; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 KEV 9

DB 4 KEV 6

RESULT 6  
157532  
gene Thielow protein - rat (fragment)  
C/Species: Rattus sp. (rat)  
C/Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 28-Feb-1997  
C/Accession: I57532  
R/Banerjee-Basu, S.; Buonanno, A.  
Mol. Cell. Biol. 13, 7019-7028, 1993  
A/Title: cis-acting sequences of the rat tropomyosin I slow gene confer tissue- and develop  
A/Reference number: I57532; MUID:94019373; PMID:8413291  
A/Accession: I57532  
A/Status: Preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-8 <RES>  
A/Cross-references: GB:S66172; NID:G432603  
C/Genetics:  
A/Genes: Thielow

Query Match 31.0%; Score 13; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 1 MPEV 4

RESULT 7  
A39308  
glycine reductase (EC 1.4.99.-) sulphydryl protein C, alpha chain - Clostridium sticklandii  
C/Species: Clostridium sticklandii  
C/Date: 19-Jun-1992 #sequence\_revision 19-Jun-1992 #text\_change 09-Jul-2004  
C/Accession: A39308  
R/Stadman, T.C.; Davis, J.N.  
J. Biol. Chem. 266, 22147-22153, 1991  
A/Title: Glycine reductase protein C. Properties and characterization of its role in th  
A/Reference number: A39308; MUID:92042141; PMID:1939235  
A/Accession: A39308  
A/Status: Preliminary  
A/Molecule type: protein  
A/Residues: 1-8 <STA>  
A/Cross-references: UNIPROT:Q7M0L0  
C/Function:  
A/Description: glycine reductase complex catalyzes the reductive deamination of glycine  
C/Keywords: ATP; oxidoreductase

Query Match 31.0%; Score 13; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 1 MKPV 4

RESULT 8  
S13889  
phosphoenolpyruvate carboxylase (EC 4.1.1.31) - maize  
C/Species: Zea mays (maize)  
C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C/Accession: S13889  
R/Jiao, J.; Chollet, R.  
Arch. Biochem. Biophys. 283, 300-305, 1990  
A/Title: Regulatory phosphorylation of serine-15 in maize phosphoenolpyruvate carboxyla  
A/Reference number: S13889; MUID:91112741; PMID:2148863  
A/Accession: S13889  
A/Status: Preliminary  
A/Molecule type: protein  
A/Residues: 1-9 <JIA>



A:Cross-references: UNIPROT:Q43267; UNIPROT:Q41197  
A:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 31.0%; Score 13; DB 2; Length 9;  
Best Local Similarity 28.6%; Pred. No. 2.8e+05;  
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANK 7  
| : :  
Db 3 SIDAQLR 9

RESULT 9  
S43014  
hypothetical protein URF-2Y - Yersinia enterocolitica transposon TN3926  
C:Species: Yersinia enterocolitica  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
C:Accession: S43014

R:Osbourn, S.E.V.; Turner, A.K.; Grinstead, J.  
submitted to the EMBL Data Library, March 1994

A:Description: The structure of the bacterial transposable element, TN3926.  
A:Reference number: S43011

A:Accession: S43014  
A:Molecule type: DNA

A:Residues: 1-4 <OSB>  
A:Cross-references: EMBL:X78059; NID:g460067; PIDN:CAAC4980.1; PID:9581836

C:Keywords: mobile element; transposon TN3926

Query Match 28.6%; Score 12; DB 2; Length 4;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LRAN 5  
| : :  
Db 1 NNAN 4

RESULT 10

PT0295  
Ig heavy chain CRD3 region (clone 5-91) - human (fragment)

C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C:Accession: PT0295  
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shale, S.; Caton, A.J.; Rivera, G.  
J. Exp. Med. 173:395-407, 1991

A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J  
A:Reference number: PT0222; MUID:91108337; PMID:1899102

A:Accession: PT0295

A:Molecule type: DNA  
A:Residues: 1-5 <YAM>

A:Experimental source: B lymphocyte  
C:Keywords: heterotrimer; immunoglobulin

Query Match 28.6%; Score 12; DB 2; Length 5;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTAN 5  
| : :  
Db 1 SKTRN 5

RESULT 11

C56793  
platelet glycoprotein GPIIa - human (fragment)

C:Species: Homo sapiens (man)  
C:Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 07-Feb-1997

C:Accession: C56793  
R:Catimel, B.; Parenttler, S.; Leung, L.L.; McGregor, J.L.  
Biochem. J. 279, 419-425, 1991

A:Title: Separation of important new platelet glycoproteins (GPIa, GPIc, GPIc\*, GPIIa and  
A:Reference number: A56793; MUID:92061944; PMID:1953640

A:Accession: C56793  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <CAT>  
A:Experimental source: platelet  
C:Keywords: glycoprotein

Query Match 28.6%; Score 12; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NMKE 8  
| : :  
Db 2 NLDE 5

RESULT 12

A39081  
amine oxidase (copper-containing) (EC 1.4.3.6) - yeast (Pichia angusta) (fragment)

C:Species: Pichia angusta  
C:Date: 31-Dec-1993 #sequence\_revision 03-Feb-1994 #text\_change 20-Apr-2000

C:Accession: A39081  
R:Mu, D.; Jones, S.M.; Smith, A.J.; Brown, D.E.; Dooley, D.M.; Kilman, J.P.  
J. Biol. Chem. 267, 7979-7982, 1992

A:Title: Tyrosine codon corresponds to topa quinone at the active site of copper amine  
A:Reference number: A39081; MUID:92235001; PMID:1569055

A:Accession: A39081

A:Molecule type: protein  
A:Residues: 1-7 <MOA>  
C:Keywords: copper; oxidoreductase; quinoprotein; topaquinone  
F:4/Modified site: topaquinone (Tyr) #status experimental

Query Match 28.6%; Score 12; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 ANKEV 9  
| : :  
Db 2 ANEYV 7

RESULT 13

S45648  
Na<sup>+</sup>-transporting ATP synthase (EC 3.6.1.-) alpha chain - Acetobacterium woodii (fragment)

C:Species: Acetobacterium woodii  
C:Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 07-May-1999

C:Accession: S45648  
R:Reidinger, J.; Mueller, V.  
Eur. J. Biochem. 223, 275-283, 1994

A:Title: Purification of ATP synthase from Acetobacterium woodii and identification as  
A:Reference number: S45648; MUID:94307271; PMID:8033902

A:Accession: S45648

A:Molecule type: protein  
A:Residues: 1-3/4-7 <RET>

A:Experimental source: DSM 1030  
C:Keywords: hydrolase

Query Match 28.6%; Score 12; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NMKE 8  
| : :  
Db 2 NLDE 5

RESULT 14

S59902  
glutathione transferase (EC 2.5.1.18) P - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)  
C:Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999

C:Accession: S59902

R:Nishihira, J.; Hibiy, Y.; Sakai, M.; Nishi, S.; Kumazaki, T.; Ohki, S.; Sakamoto, W.  
 Biochim. Biophys. Acta 1252, 233-238, 1995  
 A>Title: The C-terminal region, Arg(201)-Gln(209), of glutathione transferase P contribu  
 A:Reference number: S59902; MUID:96049505; PMID:7578228  
 A:Accession: S59902  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-9 <NIS>  
 C:Superfamily: glutathione transferase  
 C:Keywords: transferase

Query Match 28.6%; Score 12; DB 2; Length 9;  
 Best local similarity 28.6%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 LTANKK 8  
 : | :  
 DB 3 INNGKKQ 9

## RESULT 15

T46627

hypotheical protein c4 - lobiolly pine

C:Species: Pinus taeda (lobiolly pine)

C&gt;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000

C:Accession: T46627

R:Chang, S.; Purysa, J.; Funkhouser, E.A.; Newton, R.J.; Cairney, J.

Submitted to the EMBL Data Library, July 1995

A:Description: Cloning of a chitinase homolog which lacks chitin binding sites and is do

A:Reference number: 223105

A:Accession: T46627

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-4 &lt;CHA&gt;

A:Cross-references: EMBL:U11309; NID:G974285; PID:G974282

A:Experimental source: Strain 86PT2x86PT3; 8 month seedlings

## Query Match

Best local similarity 26.2%; Score 11; DB 2; Length 4;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 MKEV 9  
 : | :  
 DB 1 MKLV 4

Search completed: November 30, 2004, 08:35:17  
 Job time : 51 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 30, 2004, 08:16:03 ; Search time 58 Seconds  
(without alignments)  
89,282 Million cell updates/sec

Title: US-10-008-377a-7  
Perfect score: 42  
Sequence: 1 SLTANKKEV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1598

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_02:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	45.2	8	1	LPMS STAP
2	13	45.2	8	2	Q7DKL7
3	13	45.2	8	2	BAA34538
4	17	40.5	8	2	Q9AGP4
5	16	38.1	9	2	Q78337
6	16	38.1	9	2	Q81964
7	16	38.1	9	2	Q81966
8	16	38.1	9	2	Q81968
9	16	38.1	9	2	Q7F8S7
10	16	38.1	9	2	Q8BR43
11	16	38.1	9	2	Q8BR44
12	16	38.1	9	2	Q8BR45
13	16	38.1	9	2	Q8BR48
14	16	38.1	9	2	Q8BRK0
15	16	38.1	9	2	Q8HSL0
16	16	38.1	9	2	Q8HSL5
17	16	38.1	9	2	Q8MDU2
18	16	38.1	9	2	Q8ME56
19	16	38.1	9	2	Q8ME58
20	16	38.1	9	2	Q8T387
21	16	38.1	9	2	Q8T389
22	16	38.1	9	2	Q8T385
23	16	38.1	9	2	Q8T388
24	16	38.1	9	2	Q8T389
25	16	38.1	9	2	Q8W779
26	16	38.1	9	2	Q8W7U0
27	16	38.1	9	2	Q8W875
28	16	38.1	9	2	Q8W875
29	16	38.1	9	2	Q8W875
30	16	38.1	9	2	Q8W875
31	16	38.1	9	2	Q8W875

32	15	35.7	8	2	Q7G8P5
33	15	35.7	8	2	Q36898
34	15	35.7	8	2	Q799V9
35	15	35.7	9	2	Q6LB30
36	15	35.7	9	2	Q85710
37	15	35.7	9	2	Q85710
38	14	33.3	8	2	Q16428
39	14	33.3	8	2	Q16468
40	14	33.3	8	2	Q9N6M5
41	14	33.3	8	2	Q7JK31
42	14	33.3	8	2	Q7JK32
43	14	33.3	8	2	Q9WYLS
44	14	33.3	8	2	P92211
45	14	33.3	8	2	P92215

## ALIGNMENTS

## RESULT 1

LPMS STAP STANDARD; PRT; 8 AA.

AC P23211; 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 01-MAR-1992 (Rel. 21, Last annotation update)

DE Probable msa leader peptide.

OS Staphylococcus epidermidis.

OC Plasmid pUL5050.

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI\_Taxid=1282;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=91041730; PubMed=2233255;

RA ROSE J.I., Eady E.A., Cove J.H., Cunliffe W.J., Baumberg S.,

RT WOOLTON J.C.; "Inducible erythromycin resistance in staphylococci is encoded by a

RL member of the ATP-binding transport super-gene family.";

CC Mol. Microbiol. 4:1207-1214(1990).

CC -1- FUNCTION: May regulate expression of the erythromycin resistance

CC protein.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X52085; CAA36303.1; -

DR PIR: G11157; LPSAMP.

DR Leader peptide; Plasmid.

DR SEQUENCE 8 AA; 937 MW; FA37340685BDC1A6 CRC64;

Query Match Score 19; DB 1; Length 8;  
Best Local Similarity 50.0%; Pred. No. 1,8e+06;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LTANKK 7

DB 1 MTRSMR 6

RESULT 2

Q7DKL7 PRELIMINARY; PRT; 8 AA.

AC Q7DKL7; 05-JUL-2004 (TREMBLrel. 27, Created)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)

DE Leader peptide.

OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OK NCBI\_TaxID=1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MS968;  
 RX MEDLINE=99026951; PubMed=9809423;  
 RA Matsuo M., Endou K., Kobayashi H., Inoue M., Nakajima Y.;  
 RT "A plasmid that encodes three genes for resistance to macrolide  
 antibiotics in Staphylococcus aureus";  
 RL FEMS Microbiol. Lett. 167:221-227(1998).  
 DR EMBL, AB013298; BAA34538.1; FA37340685BDC1A6 CRC64;  
 SQ SEQUENCE 8 AA; 937 MW; FA37340685BDC1A6 CRC64;

Query Match 45.2%; Score 19; DB 2; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LTANMK 7  
 DB 1 MTASMR 6

RESULT 3  
 ID BAA34538 PRELIMINARY; PRT; 8 AA.  
 AC BAA34538;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE Leader peptide.  
 OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 CC Bacillus/Staphylococcus group; Staphylococcus.  
 OK NCBI\_TaxID=1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MS968;  
 RA Matsuo M., Endou K., Kobayashi H., Inoue M., Nakajima Y.;  
 RT "A plasmid that encodes three genes for resistance to macrolide  
 antibiotics in Staphylococcus aureus";  
 RL FEMS Microbiol. Lett. 167:221-227(1998).  
 DR EMBL, AB013298; BAA34538.1;  
 SQ SEQUENCE 8 AA; 937 MW; FA37340685BDC1A6 CRC64;

Query Match 45.2%; Score 19; DB 2; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LTANMK 7  
 DB 1 MTASMR 6

RESULT 4  
 ID Q9AGP4 PRELIMINARY; PRT; 8 AA.  
 AC Q9AGP4;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Serine hydroxymethyltransferase (Fragment).  
 GN Name=g1YA;  
 OS Arabidopsis sp. 11N.  
 OC Bacteria; Actinobacteriales; Actinobacteridae; Actinomycetales;  
 CC Micrococcales; Micrococcaceae; Arthrobacter.  
 OK NCBI\_TaxID=153502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=11N;  
 RA Meskys R., Harris R.J., Casalte V., Baerian J., Scritton N.S.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AF329478; AAK16486.1;

DR GO; GO:0008168; F:methyltransferase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 KM Methyltransferase; Transferase.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 898 MW; 6B1870533372457 CRC64;

Query Match 40.5%; Score 17; DB 2; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 NMKEV 9  
 DB 2 NUKKI 6

RESULT 5  
 ID Q78337 PRELIMINARY; PRT; 9 AA.  
 AC Q78337; 082778;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit (rbcl  
 (Fragment)).  
 GN Name=rbcl;  
 OS Caloglossa lepreurii.  
 CC Chloroplast.  
 OC Eukaryota; Rhodophyta; Floridophyceae; Ceramiales; Delesseriaceae;  
 OK NCBI\_TaxID=76903;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=736, 1048, 902, 490, 932, 922, 9, 880, 1053, and 1052;  
 RA Kamiya M., West J.A., King R.J., Zuccarello G.C., Tanaka J., Hara Y.;  
 RT "Evolutionary divergence in the red algae Caloglossa lepreurii and C.  
 apomeioteia";  
 RL J. Phycol. 34:361-370(1998).

[2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=736, 1048, 902, 490, 932, 922, 9, 880, 1053, and 1052;  
 RA Kamiya M., Tanaka J., King R.J., West J.A., Zuccarello G.C., Kawai H.;  
 RT "Reproductive and genetic distinction between broad and narrow  
 ecotypes of Caloglossa contigua (Delesseriaceae, Rhodophyta).";  
 RL J. Phycol. 36:356-367(1999).

[3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Virginia USA, and Georgia USA;  
 RA Kamiya M., West J.A., Zuccarello G.C., Kawai H.;  
 RT "Caloglossa intermedia, sp. nov. (Rhodophyta) from the western  
 atlantic coast: molecular and morphological analyses with special  
 reference to C. lepreurii and C. monosticha.";

[4]  
 RP SEQUENCE FROM N.A.  
 RA Zuccarello G.C., Bartlett J., Yeates P.H.;  
 RT "Differentiation of Caloglossa lepreurii (Rhodophyta) populations in  
 Eur. J. Phycol. 0:0-0(2000)."  
 [5]  
 RP SEQUENCE FROM N.A.  
 RA West J.A., Zuccarello G.C., Kamiya M.;  
 RT "Reproductive patterns of Caloglossa species (Delesseriaceae,  
 Rhodophyta) from Australia and New Zealand: multiple origins of  
 asexuality in C. lepreurii. Literature review on asexual, mixed-  
 phase, bisexual and sexual compatibility.";  
 RL Phycol. Res. 49:183-200(2001).  
 DR EMBL, D87813; BAA31279.1;  
 DR EMBL, D89959; BAA31303.1;  
 DR EMBL, AB030267; BAA90255.1;  
 DR EMBL, AB030268; BAA90257.1;  
 DR EMBL, D89949; BAA31283.1;  
 DR EMBL, D89951; BAA31287.1;  
 DR EMBL, D89952; BAA31289.1;



```

DE Ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit
DE (Fragment).
GN Name=tdcl;
OC Caloglossa ogasawarensis.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Delesseriaceae;
OC Caloglossa.
OX NCBI_TaxID=76907;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=596;
RA Kamiya M., West J.A., King R.J., Zuccarello G.C., Tanaka J., Hara Y.;
RT "Evolutionary divergence in the red algae Caloglossa lepreurii and C.
RT apomeiotica."
RL Phycologia 34:361-370(1998).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=596;
RA Kamiya M., Tanaka J., King R.J., West J.A., Zuccarello G.C., Kawai H.;
RT "Reproductive patterns of Caloglossa species (Delesseriaceae,
RT Rhodophyta) from Australia and New Zealand: multiple origins of
RT asexuality in C. lepreurii. Literature review on apomixis, mixed-
RT phase, bisexuality and sexual compatibility."
RL Phycol. Res. 49:183-200(2001).
DR EMBL; D89961; EAA3307.1; -.
DR EMBL; AF340182; AAK98091.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 TANM 6
DB 6 TANV 9

```

```

RESULT 9
ID Q76FS7 PRELIMINARY; PRT; 9 AA.
AC Q76FS7;
DT 05-JUL-2004 (TRENBLREL. 27, Created)
DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)
DE Rubisco large subunit (Fragment).
GN Name=tdcl;
OS Hypnea japonica.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Gigartinales; Hypneaceae;
OC Hypnea.
OX NCBI_TaxID=105606;
RN [1]
RN SEQUENCE FROM N.A.
RA Yano T., Kamiya M., Arai S., Kawai H.;
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB104706; BAC82415.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 TANM 6
DB 6 TANV 9

```

```

QY 3 TANM 6
DB 6 TANV 9

RESULT 10
ID Q8HB43 PRELIMINARY; PRT; 9 AA.
AC Q8HB43;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)
DE Ribulose 1,5-bisphosphate large subunit (Fragment).
GN Name=tdcl;
OS Gracilaria perplexa.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Gracilariiales; Gracilariaceae;
OC Gracilaria.
OX NCBI_TaxID=197869;
RN [1]
RN SEQUENCE FROM N.A.
RA Byrne K., Zuccarello G.C., West J.A., Liao M.-L., Kraft G.;
RT "Gracilaria species (Gracilariaceae, Rhodophyta) from southeastern
RT Australia, including a new species, G. perplexa sp. nov.: Morphology,
RT molecular relationships and agar content."
RL Phycol. Res. 50:295-311(2002).
DR EMBL; AY131305; AAN07038.1; -.
DR EMBL; AY131307; AAN07040.1; -.
DR EMBL; AY131308; AAN07042.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 TANM 6
DB 6 TANV 9

```

```

RESULT 11
ID Q8HB44 PRELIMINARY; PRT; 9 AA.
AC Q8HB44;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)
DE Ribulose 1,5-bisphosphate large subunit (Fragment).
GN Name=tdcl;
OS Gracilaria cliftonii.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Gracilariiales; Gracilariaceae;
OC Gracilaria.
OX NCBI_TaxID=206548;
RN [1]
RN SEQUENCE FROM N.A.
RA Byrne K., Zuccarello G.C., West J.A., Liao M.-L., Kraft G.;
RT "Gracilaria species (Gracilariaceae, Rhodophyta) from southeastern
RT Australia, including a new species, G. perplexa sp. nov.: Morphology,
RT molecular relationships and agar content."
RL Phycol. Res. 50:295-311(2002).
DR EMBL; AY131302; AAN07030.1; -.
DR EMBL; AY131303; AAN07032.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;

```

Best Local Similarity 75.0%; Pred. No. 1.8e-06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TANM 6  
DB 6 TANV 9

## RESULT 12

OSHRB45 PRELIMINARY; PRT; 9 AA.

AC OSHR45:  
DT 01-MAR-2003 (TREMURel. 23, Created)  
DT 01-MAR-2003 (TREMURel. 23, Last sequence update)  
DT 05-JUN-2004 (TREMURel. 27, Last annotation update)  
DE Ribulose-1,5-bisphosphate large subunit (Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit) (Fragment).  
GN Name=dbcl;  
OS Gracilaria chilensis (Red alga).  
OC Chloroplast.  
OC Eukaryota; Rhodophyta; Florideophyceae; Gracilariaceae; Gracilariaceae;  
OC Gracilariaceae.  
OK NCBI\_TaxID=2775;  
RN 11

SEQUENCE FROM N.A.  
RA Byrne K., Zuccarello G.C., West J.A., Liao M.-L., Kraft G.,  
RT "Gracilaria species (Gracilariaceae, Rhodophyta) from southeastern Australia, including a new species, G. perplexa sp. nov.: Morphology, molecular relationships and agar content."  
RL Phycol. Res. 50:295-311(2002).  
RN 12

SEQUENCE FROM N.A.  
RA Cohen S., Fageron S., Martinez E., Correa J.A., Viard F.,  
RA Destombe C., Valero M.,  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY131299; AN07024.1; -  
DR EMBL: AY131300; AN07026.1; -  
DR EMBL: AY131301; AN07028.1; -  
DR EMBL: AY423840; AA84581.1; -  
DR GO:GO:0009507; C:chloroplast; IEA.  
KM Chloroplast.  
FT NON\_TER  
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.8e-06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TANM 6  
DB 6 TANV 9

## RESULT 13

OSHRJ8 PRELIMINARY; PRT; 9 AA.

ID OSHRJ8:  
AC OSHRJ8:  
DT 01-MAR-2003 (TREMURel. 23, Created)  
DT 01-MAR-2003 (TREMURel. 23, Last sequence update)  
DT 01-JUN-2003 (TREMURel. 24, Last annotation update)  
DE Ribulose-1,5-bisphosphate large subunit (Fragment).  
GN Name=dbcl;  
OS Ptilophora prolifera.  
OC Chloroplast.  
OC Eukaryota; Rhodophyta; Florideophyceae; Gelidiales; Gelidiales;  
OC Ptilophora.  
OK NCBI\_TaxID=143014;  
RN 11

SEQUENCE FROM N.A.  
RA Byrne K., Zuccarello G.C., West J.A., Liao M.-L., Kraft G.,  
RT "Gracilaria species (Gracilariaceae, Rhodophyta) from southeastern Australia, including a new species, G. perplexa sp. nov.: Morphology, molecular relationships and agar content."  
RL Phycol. Res. 50:295-311(2002).  
RN 12

RL Phycol. Res. 50:295-311(2002).

DR EMBL: AY131312; AN07050.1; -  
DR GO:GO:0009507; C:chloroplast; IEA.  
KM Chloroplast.  
FT NON\_TER  
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.8e-06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TANM 6  
DB 6 TANV 9

## RESULT 14

OSHRK0 PRELIMINARY; PRT; 9 AA.

AC OSHRKO:  
DT 01-MAR-2003 (TREMURel. 23, Created)  
DT 01-MAR-2003 (TREMURel. 23, Last sequence update)  
DT 01-JUN-2003 (TREMURel. 24, Last annotation update)  
DE Ribulose-1,5-bisphosphate large subunit (Fragment).  
GN Name=dbcl;  
OS Gracilaria secundata.  
OC Chloroplast.  
OC Eukaryota; Rhodophyta; Florideophyceae; Gracilariaceae; Gracilariaceae;  
OC Gracilariaceae.  
OK NCBI\_TaxID=172973;  
RN 11

SEQUENCE FROM N.A.  
RA Byrne K., Zuccarello G.C., West J.A., Liao M.-L., Kraft G.,  
RT "Gracilaria species (Gracilariaceae, Rhodophyta) from southeastern Australia, including a new species, G. perplexa sp. nov.: Morphology, molecular relationships and agar content."  
RL Phycol. Res. 50:295-311(2002).  
DR EMBL: AY131311; AN07048.1; -  
DR GO:GO:0009507; C:chloroplast; IEA.  
KM Chloroplast.  
FT NON\_TER  
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.8e-06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TANM 6  
DB 6 TANV 9

## RESULT 15

OSHSLO PRELIMINARY; PRT; 9 AA.

ID OSHSLO:  
AC OSHSLO:  
DT 01-MAR-2003 (TREMURel. 23, Created)  
DT 01-MAR-2003 (TREMURel. 23, Last sequence update)  
DT 01-JUN-2003 (TREMURel. 24, Last annotation update)  
DE Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (Fragment).  
GN Name=dbcl;  
OS Murayella pericladus.  
OC Chloroplast.  
OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Rhodomeleaceae;  
OC Murayella.  
OK NCBI\_TaxID=110473;  
RN 11

SEQUENCE FROM N.A.  
RA Zuccarello G.C., Sanderoock B., West J.A.,  
RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF458764; AA014538.1; -  
DR GO:GO:0009507; C:chloroplast; IEA.

KW Chloroplast. 1 1  
 FT NON TER  
 SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B71AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 TANN 6  
 |||:  
 Db 6 TANN 9

Search completed: November 30, 2004, 08:34:16  
 Job time : 60 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 08:14:33 / Search time 114 Seconds  
(without alignments)  
28.321 Million cell updates/sec

Title: US-10-008-377a-7

Perfect score: 42

Sequence: 1 SLTANKKEV 9

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 313949

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: A\_Geneseq\_23Sep04:\*

1: geneseqp1980s:\*\n2: geneseqp1980s:\*\n3: geneseqp2001s:\*\n4: geneseqp2002s:\*\n5: geneseqp2003as:\*\n6: geneseqp2003bs:\*\n7: geneseqp2003bs:\*\n8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	42	100.0	9 6 AAO22989	Aao22989 Human p62
2	24	57.1	9 4 AAB07594	Aab07594 Human FUM
3	24	57.1	9 6 ABR23168	AbR23168 Human can
4	24	57.1	9 6 ABR21768	AbR21768 Human can
5	24	57.1	9 6 ABR20369	AbR20369 Human can
6	23	54.8	9 6 AB019931	Ab019931 MHC bindi
7	23	54.8	9 6 AB020118	Ab020118 MHC bindi
8	23	54.8	9 6 ABR23049	AbR23049 Human can
9	23	54.8	9 6 ABR20246	AbR20246 Human can
10	23	54.8	9 6 ABR21637	AbR21637 Human can
11	22	52.4	9 6 ADH68371	AdH68371 RSV F-pro
12	22	52.4	9 4 AAB46277	Aab46277 HPV type
13	22	52.4	9 8 ADH40759	AdH40759 Human CD4
14	22	52.4	9 8 ADH40364	AdH40364 Human CD4
15	21	50.0	9 8 AAG60252	Aag60252 HOP prote
16	21	50.0	9 6 ABM66287	Abm66287 Prolonib
17	20	47.6	9 6 ADH68370	AdH68370 RSV F-pro
18	20	47.6	7 3 AAY81479	Aay81479 Murine me
19	20	47.6	7 4 AAB98536	Aab98536 Human pep
20	20	47.6	8 3 AAB21665	Aab21665 Immunomod
21	20	47.6	8 3 AAB09862	Aab09862 Immunomod
22	20	47.6	8 5 AAB20699	Aab20699 Immunomod
23	20	47.6	8 7 ABR83118	AbR83118 Beta-gala
24	20	47.6	9 2 AAB57750	Aab57750 Peptide (
25	20	47.6	9 2 AAY55417	Aay55417 HLA bindi

26	20	47.6	9 3 AAB09858	Aab09858 Immunomod
27	20	47.6	9 4 AAB75636	Aab75636 HLA class
28	20	47.6	9 5 AAB20695	Aab20695 Immunomod
29	20	47.6	9 5 AAY94681	Aay94681 Human nov
30	20	47.6	9 5 AAY94477	Aay94477 Human nov
31	20	47.6	9 5 AAY94091	Aay94091 Human nov
32	20	47.6	9 6 ABR22026	AbR22026 Human can
33	20	47.6	9 6 ABR23439	AbR23439 Human can
34	20	47.6	9 6 ABR23216	AbR23216 Human can
35	20	47.6	9 6 ABR20255	AbR20255 Human can
36	20	47.6	9 6 ABR21647	AbR21647 Human can
37	20	47.6	9 6 ABR21809	AbR21809 Human can
38	20	47.6	9 6 ABR20636	AbR20636 Human can
39	20	47.6	9 6 ABR23063	AbR23063 Human can
40	20	47.6	9 6 ABR20415	AbR20415 Human can
41	20	47.6	9 6 ADK10387	AdK10387 Human pap
42	19	45.2	6 2 AAW83819	Aaw83819 Peptide s
43	19	45.2	6 2 AAW83709	Aaw83709 Peptide s
44	19	45.2	7 5 ABA49017	AbA49017 Zinc fing
45	19	45.2	7 6 ABP70385	AbP70385 Tryptic p

## ALIGNMENTS

RESULT 1  
AAO22989 standard; peptide; 9 AA.  
ID AAO22989; (first entry)  
AC AAO22989;  
XX 17-SEP-2003  
DT Human p621-629 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.  
DE HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
XX cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;  
XX oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
XX t(2; 5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
XX p621-629.  
OS Homo sapiens.  
XX WO2003042243-A2.  
XX 22-MAY-2003.  
XX 14-NOV-2002; 2002WO-EP012764.  
XX 15-NOV-2001; 2001US-00008377.  
XX (NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.  
XX (SUDD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX Gambacorti-Passerini C, Passoni L;  
XX WPI; 2003-441791/41.  
XX New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
XX useful for preparing a composition for treating ALK-positive lymphoma,  
XX neuroblastoma or ALK-expressing neoplasia.  
XX Claim 1; Page 7; 33pp; English.  
XX The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma  
XX kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a  
XX cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase  
XX which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase  
XX fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell  
XX lines. More than 50% of ALCL cases possess a t(2;5) chromosomal  
XX translocation that leads to the expression of the NPM/ALK fusion protein  
XX which forms a potent oncogene when constitutively activated. Translocated  
XX ALK is a widely expressed tumour-associated antigen characteristic of ALK

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p621-629  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 XX

SO Sequence 9 AA;

Query Match 100.0%; Score 42; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLTANKEV 9  
 |||||  
 1 SLTANKEV 9

RESULT 2  
 AAE07594  
 ID AAE07594 standard; peptide; 9 AA.

AC AAE07594;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX

DE Human PUMP-1 peptide (residues 54-62).

KW PUMP-1 protease; cancer; neoplastic state; malignancy; ovary; lung;  
 KM prostate; colon; cytostatic; gene therapy; vaccine; immunogen; human.

XX Homo sapiens.

OS WO200154712-A1.

PN 02-AUG-2001.

PD 26-JUN-2001; 2001WO-US002698.

PR 27-JAN-2000; 2000US-00492543.

PA (UYAR-) UNIV ARKANSAS.

PI O'Brien TJ;

XX WPI; 2001-496835/54.

PT Diagnosing cancer in an individual, useful for early detection of ovarian  
 PT cancer, lung cancer or prostate cancer, comprises determining the  
 PT presence of PUMP-1 protease in biological sample obtained from  
 PT individual.

PS Example 20; Page 65; 145pp; English.

XX The patent discloses compositions and methods for the early diagnosis of  
 CC ovarian cancer. The method involves obtaining a biological sample from an  
 CC individual and detecting PUMP-1 protease in the sample, where the  
 CC presence of PUMP-1 in the sample is indicative of the presence of cancer  
 CC in the individual and the absence of PUMP-1 is indicative of the absence  
 CC of cancer in the individual. The method is useful for the early detection  
 CC or diagnosis of ovarian cancer and other neoplastic state or malignancies  
 CC (e.g. lung cancer, prostate cancer, colon cancer or other cancers in  
 CC which PUMP-1 is overexpressed). It is also useful for diagnosing whether  
 CC an individual has cancer, is suspected of having cancer or is at risk of  
 CC getting cancer. The method is also used to detect and treat malignant  
 CC hyperplasia. The PUMP-1 proteins are also useful for vaccinating against  
 CC neoplastic states. The oligonucleotide is useful for treating neoplastic  
 CC states such as lung cancer, prostate cancer, colon cancer or other  
 CC cancers in which PUMP-1 is overexpressed. The present sequence is PUMP-1  
 CC peptide (residues 54-62) from human. This peptide is a putative immunogen

CC and binds strongly to HLA A1 and is used to inoculate an individual  
 CC against PUMP-1

SO Sequence 9 AA;

Query Match 57.1%; Score 24; DB 4; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SLTANKEV 9  
 |||||  
 1 SLTANKEV 9

RESULT 3  
 ABR23168  
 ID ABR23168 standard; peptide; 9 AA.

AC ABR23168;

XX 19-MAY-2003 (first entry)

DE Human cancer-related protein 185P2C9 HLA peptide #3003.

KW Human; cytostatic; vaccine; cancer; immune response; HLA;  
 KM human leukocyte antigen.

XX Homo sapiens.

OS WO200283921-A2.

PN 24-OCT-2002.

PD 10-APR-2002; 2002WO-US011654.

PR 10-APR-2001; 2001US-0282739P.

PR 10-APR-2001; 2001US-0283112P.

PR 25-APR-2001; 2001US-028630P.

PA (AGEN-) AGENSYS INC.

PI Jakobovits A, Challita-Bid PM, Faris M, Ge W, Hubert RS;

PT Morrison K, Morrison RK, Raitano AB;

PT New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides, useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.

PS Claim 13; Page 354; 1021pp; English.

XX The present invention relates to novel human cancer-related genes and  
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention

SO Sequence 9 AA;

Query Match 57.1%; Score 24; DB 6; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SLTANKEV 9

Db 1 NLSDMKREV 9

RESULT 4  
ABR21768  
ID ABR21768 standard; peptide; 9 AA.  
XX  
AC ABR21768;  
XX  
DT 19-MAY-2003 (first entry)  
XX  
DE Human cancer-related protein 185P2C9 HLA peptide #1603.  
XX  
KW Human; cytostatic; vaccine; cancer; immune response; HLA;  
KW human leukocyte antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO200283921-A2.  
XX  
PD 24-OCT-2002.  
XX  
PF 10-APR-2002; 2002WO-US011654.  
XX  
PR 10-APR-2001; 2001US-0282739P.  
PR 10-APR-2001; 2001US-0283112P.  
PR 25-APR-2001; 2001US-0286630P.  
XX  
PA (AGEN-) AGENSYS INC.  
XX  
PI Jakobovits A, Chailita-Bid PM, Paris M, Ge W, Hubert RS;  
PI Morrison K, Morrison RK, Raitano AB;  
XX  
DR WPI; 2003-075555/07.  
XX  
PT New composition comprising a substance that modulates the structure of  
PT proteins and polynucleotides, useful for therapeutic, prognostic and  
PT diagnostic reagents for eliciting cellular or humoral immune response in  
PT cancer patients.  
XX  
PS Claim 13; Page 338; 1021pp; English.  
XX  
CC The present invention relates to novel human cancer-related genes and  
CC proteins (ABR278120-ABR278168 and ABR01789-ABR01861). The genes and  
CC proteins are useful for eliciting a humoral or cellular immune response.  
CC The genes are useful as probes and primers for the amplification and/or  
CC detection of genes, mRNAs or their fragments, as reagents for the  
CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
CC directing the expression of the protein, as tools for modulating or  
CC inhibiting the expression of genes and/or translation of transcripts, and  
CC as therapeutic agents. The proteins and peptides are useful as  
CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
CC from the invention  
XX  
SQ Sequence 9 AA;  
XX

Query Match 57.1%; Score 24; DB 6; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
: : : : :  
1 NLSDMKREV 9

Db 1 NLSDMKREV 9

RESULT 5  
ABR20369  
ID ABR20369 standard; peptide; 9 AA.  
XX  
AC ABR20369;  
XX

DT 19-MAY-2003 (first entry)  
XX  
DE Human cancer-related protein 185P2C9 HLA peptide #204.  
XX  
KW Human; cytostatic; vaccine; cancer; immune response; HLA;  
KW human leukocyte antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO200283921-A2.  
XX  
PD 24-OCT-2002.  
XX  
PF 10-APR-2002; 2002WO-US011654.  
XX  
PR 10-APR-2001; 2001US-0282739P.  
PR 10-APR-2001; 2001US-0283112P.  
PR 25-APR-2001; 2001US-0286630P.  
XX  
PA (AGEN-) AGENSYS INC.  
XX  
PI Jakobovits A, Chailita-Bid PM, Paris M, Ge W, Hubert RS;  
PI Morrison K, Morrison RK, Raitano AB;  
XX  
DR WPI; 2003-075555/07.  
XX  
PT New composition comprising a substance that modulates the structure of  
PT proteins and polynucleotides, useful for therapeutic, prognostic and  
PT diagnostic reagents for eliciting cellular or humoral immune response in  
PT cancer patients.  
XX  
PS Claim 13; Page 322; 1021pp; English.  
XX  
CC The present invention relates to novel human cancer-related genes and  
CC proteins (ABR278120-ABR278168 and ABR01789-ABR01861). The genes and  
CC proteins are useful for eliciting a humoral or cellular immune response.  
CC The genes are useful as probes and primers for the amplification and/or  
CC detection of genes, mRNAs or their fragments, as reagents for the  
CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
CC directing the expression of the protein, as tools for modulating or  
CC inhibiting the expression of genes and/or translation of transcripts, and  
CC as therapeutic agents. The proteins and peptides are useful as  
CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
CC from the invention  
XX  
SQ Sequence 9 AA;  
XX

Query Match 57.1%; Score 24; DB 6; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
: : : : :  
1 NLSDMKREV 9

Db 1 NLSDMKREV 9

RESULT 6  
ABU19931  
ID ABU19931 standard; peptide; 9 AA.  
XX  
AC ABU19931;  
XX  
DT 10-APR-2003 (first entry)  
XX  
DE MHC binding peptide SEQ ID No 96.  
XX  
KW Antiinflammatory; antiallergic; antiarthritic; noctropic; neuroprotective;  
KW antiinflammatory; major histocompatibility complex; MHC;  
KW autoimmune disease; T cell; B cell; allergic disease; multiple sclerosis;  
KW rheumatoid arthritis; neurodegenerative disorder; Alzheimer's disease;  
KW inflammation; gene therapy; MHC binding peptide.  
XX

OS Synthetic.  
XX WO200294981-A2.  
XX 28-NOV-2002.  
XX 16-MAY-2002; 2002WO-IL000383.  
XX 16-MAY-2001; 2001US-0280958P.  
XX 29-MAY-2001; 2001US-00865348.  
XX (TECR ) TECHNION RES & DEV FOUND LTD.  
XX Barnea E, Beer I, Ziv T, Admon A, Daessau L, Buchsbaum S;  
XX WPI; 2003-210043/20.  
XX  
XX Identifying peptides that are capable of binding to major  
XX histocompatibility complex (MHC) molecules of a particular haplotype by  
XX analyzing peptides bound to the soluble and secreted form of the MHC  
XX molecules of the particular haplotype.  
XX  
XX Claim 37; Page 169; 238pp; English.  
XX  
XX The invention relates to a novel method for identifying peptides  
XX originating from a particular cell type, which are capable of binding to  
XX major histocompatibility complex (MHC) molecules of a particular  
XX haplotype. The method comprises analyzing peptides bound to the soluble  
XX and secreted form of the MHC molecules of the particular haplotype. The  
XX method is useful for identifying peptides for treating an autoimmune  
XX disease, such as T or B cell and/or allergic disease or condition,  
XX rheumatoid arthritis, or multiple sclerosis, neurodegenerative disorders,  
XX e.g. Alzheimer's disease, or diseases associated with inflammation. The  
XX sequences of the invention may be used in a gene therapy application.  
XX This sequence represents a peptide relating to the method for identifying  
XX MHC binding peptides of the invention  
XX  
XX Sequence 9 AA;  
XX  
XX Query Match 54.8%; Score 23; DB 6; Length 9;  
XX Best Local Similarity 44.4%; Pred. No. 1.7e+06;  
XX Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX 1 SLTANMKREV 9  
XX :||:|:|:  
XX 1 ALTGHEEV 9  
XX  
XX  
XX RESULT 7  
XX ABJ20118  
XX ID ABJ20118 standard; peptide; 9 AA.  
XX  
XX AC ABJ20118;  
XX  
XX DT 10-APR-2003 (first entry)  
XX  
XX DE MHC binding peptide SEQ ID No 283.  
XX  
XX KW Antiinflammatory; major histocompatibility complex; MHC;  
XX autoinflammatory; T cell; B cell; allergic disease; multiple sclerosis;  
XX rheumatoid arthritis; neurodegenerative disorder; Alzheimer's disease;  
XX inflammation; gene therapy; MHC binding peptide.  
XX  
XX OS Synthetic.  
XX PN WO200294981-A2.  
XX PD 28-NOV-2002.  
XX PF 16-MAY-2002; 2002WO-IL000383.  
XX PT 16-MAY-2001; 2001US-0280958P.  
XX PR 29-MAY-2001; 2001US-0280958P.

PR 29-MAY-2001; 2001US-00865548.  
XX  
XX (TECR ) TECHNION RES & DEV FOUND LTD.  
XX  
XX Barnea E, Beer I, Ziv T, Admon A, Daessau L, Buchsbaum S;  
XX WPI; 2003-210043/20.  
XX  
XX Identifying peptides that are capable of binding to major  
XX histocompatibility complex (MHC) molecules of a particular haplotype by  
XX analyzing peptides bound to the soluble and secreted form of the MHC  
XX molecules of the particular haplotype.  
XX  
XX Claim 58; Page 216; 238pp; English.  
XX  
XX The invention relates to a novel method for identifying peptides  
XX originating from a particular cell type, which are capable of binding to  
XX major histocompatibility complex (MHC) molecules of a particular  
XX haplotype. The method comprises analyzing peptides bound to the soluble  
XX and secreted form of the MHC molecules of the particular haplotype. The  
XX method is useful for identifying peptides for treating an autoimmune  
XX disease, such as T or B cell and/or allergic disease or condition,  
XX rheumatoid arthritis, or multiple sclerosis, neurodegenerative disorders,  
XX e.g. Alzheimer's disease, or diseases associated with inflammation. The  
XX sequences of the invention may be used in a gene therapy application.  
XX This sequence represents a peptide relating to the method for identifying  
XX MHC binding peptides of the invention  
XX  
XX Sequence 9 AA;  
XX  
XX Query Match 54.8%; Score 23; DB 6; Length 9;  
XX Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
XX Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
XX  
XX 1 SLTANMKREV 9  
XX :||:|:|:  
XX 1 SLTANMKLEI 9  
XX  
XX  
XX RESULT 8  
XX ABR23049  
XX ID ABR23049 standard; peptide; 9 AA.  
XX  
XX AC ABR23049;  
XX  
XX DT 19-MAY-2003 (first entry)  
XX  
XX DE Human cancer-related protein 185P2C9 HLA peptide #2884.  
XX  
XX KW Human; cytostatic; vaccine; cancer; immune response; HLA;  
XX human leukocyte antigen.  
XX  
XX OS Homo sapiens.  
XX PN WO200283921-A2.  
XX PD 24-OCT-2002.  
XX PF 10-APR-2002; 2002WO-US011654.  
XX PR 10-APR-2001; 2001US-0282739P.  
XX PR 10-APR-2001; 2001US-0283112P.  
XX PR 25-APR-2001; 2001US-0286630P.  
XX  
XX PA (AGEN-) AGENSYS INC.  
XX PI Jakobovits A, Challita-Bid PM, Faris M, Ge W, Hubert RS;  
XX Morrison K, Morrison RK, Raitano AB;  
XX WPI; 2003-075555/07.  
XX  
XX New composition comprising a substance that modulates the structure of  
XX proteins and polynucleotides, useful for therapeutic, prognostic and  
XX

PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.  
 XX  
 PS Claim 13; Page 353; 1021pp; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention  
 CC  
 SQ Sequence 9 AA;  
 CC  
 QY Query Match 54.8%; Score 23; DB 6; Length 9;  
 DB Best Local Similarity 62.5%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 2 LTANKKEV 9  
 1 LSDMKKEV 8  
 DB  
 RESULT 9  
 ABR20246  
 ID ABR20246 standard; peptide; 9 AA.  
 XX  
 AC ABR20246;  
 XX  
 DT 19-MAY-2003 (first entry)  
 XX  
 DE Human cancer-related protein 185P2C9 HLA peptide #81.  
 XX  
 KW Human; cytosolic; vaccine; cancer; immune response; HLA;  
 KW human leukocyte antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283921-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 10-APR-2002; 2002WO-US011654.  
 XX  
 PR 10-APR-2001; 2001US-0282739P.  
 PR 10-APR-2001; 2001US-0283112P.  
 PR 25-APR-2001; 2001US-0286630P.  
 XX  
 PA (AGEN-) AGENSYS INC.  
 XX  
 PI Jakovcic A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 XX  
 DR WPI; 2003-075555/07.  
 XX  
 PT New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides; useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.  
 XX  
 PS Claim 13; Page 321; 1021pp; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the

CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention  
 CC  
 SQ Sequence 9 AA;  
 CC  
 QY Query Match 54.8%; Score 23; DB 6; Length 9;  
 DB Best Local Similarity 62.5%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 2 LTANKKEV 9  
 1 LSDMKKEV 8  
 DB  
 RESULT 10  
 ABR21637  
 ID ABR21637 standard; peptide; 9 AA.  
 XX  
 AC ABR21637;  
 XX  
 DT 19-MAY-2003 (first entry)  
 XX  
 DE Human cancer-related protein 185P2C9 HLA peptide #1472.  
 XX  
 KW Human; cytosolic; vaccine; cancer; immune response; HLA;  
 KW human leukocyte antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283921-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 10-APR-2002; 2002WO-US011654.  
 XX  
 PR 10-APR-2001; 2001US-0282739P.  
 PR 10-APR-2001; 2001US-0283112P.  
 PR 25-APR-2001; 2001US-0286630P.  
 XX  
 PA (AGEN-) AGENSYS INC.  
 XX  
 PI Jakovcic A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 XX  
 DR WPI; 2003-075555/07.  
 XX  
 PT New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides; useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.  
 XX  
 PS Claim 13; Page 337; 1021pp; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention  
 CC  
 SQ Sequence 9 AA;

Query Match 54.8%; Score 23; DB 6; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKKEV 9  
 : : : : :  
 DB 1 LSDDMKEV 8

## RESULT 11

ID ADH68371 standard; peptide; 6 AA.  
 AC ADH68371;

DE 25-MAR-2004 (first entry)

XX RSV F-protein peptide fragment #2.

XX vaccinia; metapneumovirus family; respiratory syncytial virus;  
 XX fusion protein; F protein; attenuated; stabilisation; virucide.

OS Respiratory syncytial virus.

PN WO2003097089-A2.

PD 27-NOV-2003.

PF 16-MAY-2003; 2003WO-EP005187.

XX 16-MAY-2002; 2002DE-01021836.

XX (LOHM) LOHMANN ANIMAL HEALTH GMBH & CO KG.

XX Naylor CJ;

DR WPI; 2004-012486/01.

PT Vaccine against metapneumovirus or respiratory syncytial virus, useful  
 for preventing infections in birds or humans, contains virus with  
 modified fusion protein sequence.

XX Disclosure; SEQ ID NO 78; 100pp; German.

XX This invention describes a novel vaccine against a member of the  
 CC metapneumovirus family, respiratory syncytial virus (RSV) or a virus  
 CC having significant genetic homology with them in the fusion (F) protein  
 CC region, containing a virus (or part) in which the 293-296 amino acid  
 CC region of F protein (or functional equivalents, e.g. 323-328 in RSV) has  
 CC been modified. The preferred vaccine contains a live, preferably  
 CC attenuated, virus, and is formulated with auxiliary, carrier and/or  
 CC adjuvant, optionally also with interleukins 6 (particularly for avian  
 CC metapneumovirus) or 12 or 18 (particularly for human metapneumovirus or  
 CC RSV). The modification in the F protein is a stabilisation, especially  
 CC substitution of one codon by another that requires several mutations for  
 CC reversion to the wild type, specifically by a codon with low probability  
 CC of reversion to a Glu-encoding codon. A virulent virus is modified in the  
 CC nucleic acid sequence that encodes the specified part of the F protein.  
 CC Particularly a full-length DNA copy of the genome is assembled by  
 CC ligating PCR products that include the required modification, then virus  
 CC is rescued using e.g. chickpox 17 polymerase or cellular ribosomal pool  
 CC RNA polymerase. The vaccines, also the modified F proteins, are used to  
 CC prevent disease caused by human or avian metapneumovirus or RSV.  
 CC Modification of the specified region of the F protein, which is important  
 CC for virulence, results in a stable attenuation, resistant to reversion to  
 CC the wild type.

XX Sequence 6 AA:

Query Match 52.4%; Score 22; DB 8; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 TANKKE 8  
 : : : : :  
 DB 1 TTWKE 6

RESULT 12  
 AAB46277  
 ID AAB46277 standard; peptide; 9 AA.  
 AC AAB46277;

DE 04-APR-2001 (first entry)

XX HPV type 16 L1 protein cytotoxic T-cell epitope SEQ ID NO 5.

XX Cytotoxic T cell epitope; L1 protein; antigenic; vaccine; antiviral;  
 XX antitumor; immune response; tumor.

OS Human papillomavirus.

PN DE19925199-A1.

PD 07-DEC-2000.

PF 01-JUN-1999; 99DE-01025199.

XX 01-JUN-1999; 99DE-01025199.

XX (MED1-) MEDIGENE AG.

XX Nieland J, Jochmus I;

DR WPI; 2001-072111/09.

PT New T cell epitopes from the papilloma virus L1 protein, useful for  
 detecting or inducing an immune response, e.g. in vaccines.

XX Claim 1; Page 13; 24pp; German.

XX This invention describes novel T cell epitopes (A) of the human papilloma  
 CC virus L1 protein. The invention also describes (a) compounds (I)  
 CC containing an (A) provided it is not a natural L1 protein or an  
 CC exclusively N- or C-terminal deletion variant of natural L1; (b) nucleic  
 CC acid (II) that encodes (A) or (I); (c) vector, particularly an expression  
 CC vector, that contains (II); (d) cells that contain, and preferably  
 CC present, (A); (e) complex of (A) or (I) with at least one additional  
 CC molecule (III); in vitro detection of T cell activation by a compound  
 CC containing at least one (A); (f) preparing cells of (d); and (g) test  
 CC system for in vitro detection of T cells. The products of the invention  
 CC have antiviral and antitumor activity. (A), also compounds or complexes  
 CC containing it or nucleic acid or vectors encoding it, is used for  
 CC detection of an immune response, particularly detecting L1-specific  
 CC cytotoxic T cells or to determine the L1 protein-specific antigenicity of  
 CC compounds and complexes that contain (A), and for inducing a protective  
 CC immune response against papilloma virus and related tumors, particularly  
 CC as vaccines

XX Sequence 9 AA:

Query Match 52.4%; Score 22; DB 4; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LTANKKEV 9  
 : : : : :  
 DB 1 TLQANKSEV 9

## RESULT 13

ID ADH40759 standard; peptide; 9 AA.  
 AC ADH40759;

XX ADH40759;

XX 11-MAR-2004 (first entry)  
XX Human CD42b HLA binding peptide.  
XX human; cytostatic; vaccine; SNP profile; cancer; leukaemia; HLA.  
XX Homo sapiens.  
XX WO2003106692-A2.  
XX 24-DEC-2003.  
XX 13-JUN-2003; 2003WO-EP006251.  
XX 13-JUN-2002; 2002EP-00013423.  
XX (MERCK) MERCK PATENT GMBH.  
XX Stritmatter W, Moll H;  
XX WPI; 2004-082200/08.  
XX Providing allelic variant epitope of protein based on single nucleotide  
XX polymorphism by defining target protein, screening database of protein,  
XX identifying, selecting allelic variant protein, creating variant  
XX epitopes.  
XX Disclosure; Page 96; 11pp; English.  
XX The invention relates to a novel method for providing epitopes of allelic  
XX variants of antigenic proteins from specific species based on single  
XX nucleotide polymorphism (SNP), by defining target protein/peptide or its  
XX subset, screening database of DNA encoding target protein, identifying,  
XX selecting allelic peptide/protein variants, expression product or its  
XX fragment encoded by DNA sequence having SNP, creating variant epitopes,  
XX selecting epitopes binding to MHC protein. A protein of the invention has  
XX cytostatic activity, and may have a use in a vaccine. The method is  
XX useful for generating a SNP profile of one or more individuals from a  
XX given species by applying the method for several protein from the  
XX individuals, where the SNP profile was related to disease, preferably  
XX cancer. This is useful for diagnosing a disease in an individual by  
XX generating the SNP-related polymorphic profile. A method of the invention  
XX is useful for transplanting haematopoietic stem cells from a donor to a  
XX recipient and treating cancer, preferably leukaemia, and for determining  
XX the progression, regression or onset of a treated disease. The present  
XX sequence is used in the exemplification of the invention.  
XX Sequence 9 AA;  
XX  
XX Query Match 52.4%; Score 22; DB 8; Length 9;  
XX Best Local Similarity 44.4%; Pred. No. 1.7e+06;  
XX Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
XX  
XX 1 SLTANKREV 9  
XX 1 SMTSNVASV 9  
XX  
XX RESULT 14  
XX ADH40364  
XX ADH40364 standard; peptide; 9 AA.  
XX ADH40364;  
XX 11-MAR-2004 (first entry)  
XX Human CD42b HLA class I binding peptide.  
XX human; cytostatic; vaccine; SNP profile; cancer; leukaemia; HLA.  
XX Homo sapiens.  
XX

XX WO2003106692-A2.  
XX 24-DEC-2003.  
XX 13-JUN-2003; 2003WO-EP006251.  
XX 13-JUN-2002; 2002EP-00013423.  
XX (MERCK) MERCK PATENT GMBH.  
XX Stritmatter W, Moll H;  
XX WPI; 2004-082200/08.  
XX Providing allelic variant epitope of protein based on single nucleotide  
XX polymorphism by defining target protein, screening database of protein,  
XX identifying, selecting allelic variant protein, creating variant  
XX epitopes.  
XX Disclosure; Page 88; 11pp; English.  
XX The invention relates to a novel method for providing epitopes of allelic  
XX variants of antigenic proteins from specific species based on single  
XX nucleotide polymorphism (SNP), by defining target protein/peptide or its  
XX subset, screening database of DNA encoding target protein, identifying,  
XX selecting allelic peptide/protein variants, expression product or its  
XX fragment encoded by DNA sequence having SNP, creating variant epitopes,  
XX selecting epitopes binding to MHC protein. A protein of the invention has  
XX cytostatic activity, and may have a use in a vaccine. The method is  
XX useful for generating a SNP profile of one or more individuals from a  
XX given species by applying the method for several protein from the  
XX individuals, where the SNP profile was related to disease, preferably  
XX cancer. This is useful for diagnosing a disease in an individual by  
XX generating the SNP-related polymorphic profile. A method of the invention  
XX is useful for transplanting haematopoietic stem cells from a donor to a  
XX recipient and treating cancer, preferably leukaemia, and for determining  
XX the progression, regression or onset of a treated disease. The present  
XX sequence is used in the exemplification of the invention.  
XX Sequence 9 AA;  
XX  
XX Query Match 52.4%; Score 22; DB 8; Length 9;  
XX Best Local Similarity 44.4%; Pred. No. 1.7e+06;  
XX Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
XX  
XX 1 SLTANKREV 9  
XX 1 SMTSNVASV 9  
XX  
XX RESULT 15  
XX AAG80232  
XX AAG80232 standard; peptide; 8 AA.  
XX AAG80232;  
XX 23-JAN-2002 (first entry)  
XX HOP protein TPRI domain associated oligopeptide #4.  
XX Hop protein; chaperone protein; TPR motif; tetratricic peptide repeat;  
XX three-dimensional structure; 3D; cytostatic; virucide; immunosuppressive;  
XX antiinflammatory; tumour treatment; autoimmune disease; TPR;  
XX protein coordinate data.  
XX Undetermined.  
XX DE10018335-A1.  
XX 04-OCT-2001.  
XX 13-APR-2000; 2000DE-01018335.  
XX

PR 29-MAR-2000; 2000DE-01015748.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX PI Moarefi I, Scheufler C, Hartl U, Brinker A;

XX WFI; 2002-000454/01.

XX  
PT New spatial model of polypeptide containing a TPR motif, useful for  
PT identifying inhibitors of interaction between chaperone and Hop proteins,  
PT potential therapeutics.

XX  
PS Disclosure: Fig 6; 145pp; German.

XX  
CC This invention describes a novel spatial model of a polypeptide (I)  
CC having at least one amino acid (aa) sequence (II) of a TPR (tetratricic  
CC peptide repeat) structural motif of a Hop protein, or a derivative of  
CC (II). The invention also describes a method for (1) producing a crystal  
CC containing, per asymmetric unit of its unit cell, at least one (I) and  
CC optionally another compound; (2) preparing the crystals of (1); (3)  
CC preparing compounds (III) that bind, as ligands, to a structural region  
CC of a Hop protein (or structurally related protein) by non-covalent  
CC interaction with the main and/or side chains of aa that are components of  
CC the TPR domain; (4) preparing peptides (Ia), having the spatial model of  
CC (I), that can bind to a chaperone protein (CP) but are non-functional as  
CC regards adapter function; (5) identifying compounds (A) that inhibit  
CC interaction between CP and Hop proteins; (6) (A) identified by the method  
CC of (5); (7) representing the three-dimensional (3D) structure of a  
CC polypeptide of unknown structure but containing at least one TPR motif  
CC (or its derivative or complex), using the model of (I); (8) identifying  
CC compounds (A') that inhibit interaction between CP and polypeptides  
CC containing at least one TPR motif, or its derivatives; (9) pharmaceutical  
CC compositions containing (A) or (A'); (10) preparing DNA sequences (IV)  
CC that encode a sequence that contains at least part of proteins T1 to 4,  
CC IRSF, SGT and KIA0719; (11) preparing amino acid sequences (V) encoded  
CC by (IV); and (12) preparing pharmaceutical composition containing (V),  
CC their fragments or derivatives. The products of the invention have  
CC cytostatic, virostatic, immunosuppressive and antiinflammatory activity.  
CC The models are used to identify compounds (A) that inhibit interaction  
CC between Hop and chaperone proteins, potentially useful for treating  
CC tumours and virus infections, also as immunosuppressives, for treating  
CC autoimmune diseases or other inflammatory states or guest versus host  
CC disease. This sequence represents a Hop protein TPR domain associated  
CC oligopeptide used to illustrate the method of the invention  
XX

SQ Sequence 8 AA:

Query Match 50.0%; Score 21; DB 5; Length 8;

Best Local Similarity 57.1%; Pred. No. 1.7e+06;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 TANKEV 9

DB 1 ISRMEEV 7

Search completed: November 30, 2004, 08:33:08  
Job time : 117 secs



RESULT 2  
 US-09-835-948-84  
 Sequence 84, Application US/09835948  
 Patent No. US20020160365A1  
 GENERAL INFORMATION:  
 APPLICANT: O'Brien, Timothy J.  
 TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
 TITLE OF INVENTION: Ovarian Cancer

FILE REFERENCE: D6223CIP-B/D  
CURRENT APPLICATION NUMBER: US/09/835,948  
CURRENT FILING DATE: 2001-04-16  
PRIOR APPLICATION NUMBER: US 09/492,543  
PRIOR FILING DATE: 01-27-2000  
NUMBER OF SEQ ID NOS: 189  
SEQ ID NO 84  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein  
US-09-835-948-84

Query Match  
Best Local Similarity 57.1%; Score 24; DB 9; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.4e+06;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 1 SLEAKLEM 9

RESULT 3  
US-09-835-948-141  
Sequence 141; Application US/09835948  
Patent No. US20020160365A1  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy J.  
TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
TITLE OF INVENTION: Ovarian Cancer  
FILE REFERENCE: D6223CIP-B/D  
CURRENT APPLICATION NUMBER: US/09/835,948  
CURRENT FILING DATE: 2001-04-16  
PRIOR APPLICATION NUMBER: US 09/492,543  
PRIOR FILING DATE: 01-27-2000  
NUMBER OF SEQ ID NOS: 189  
SEQ ID NO 141  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein  
US-09-835-948-141

Query Match  
Best Local Similarity 57.1%; Score 24; DB 9; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.4e+06;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 1 SLEAKLEM 9

RESULT 4  
US-10-172-597-84  
Sequence 84; Application US/10172597  
Publication No. US20030180736A1  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy J.  
TITLE OF INVENTION: Methods for the Early Diagnosis of Ovarian Cancer  
FILE REFERENCE: D6223CIP-B/D/CIP  
CURRENT APPLICATION NUMBER: US/10/172,597  
CURRENT FILING DATE: 2002-06-14  
PRIOR APPLICATION NUMBER: US 09/835,948  
PRIOR FILING DATE: 2001-04-16  
NUMBER OF SEQ ID NOS: 189  
SEQ ID NO 84  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein

US-10-172-597-84

Query Match  
Best Local Similarity 57.1%; Score 24; DB 14; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.4e+06;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 1 SLEAKLEM 9

RESULT 5  
US-10-172-597-141  
Sequence 141; Application US/10172597  
Publication No. US20030180736A1  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy J.  
TITLE OF INVENTION: Methods for the Early Diagnosis of Ovarian Cancer  
FILE REFERENCE: D6223CIP-B/D/CIP  
CURRENT APPLICATION NUMBER: US/10/172,597  
CURRENT FILING DATE: 2002-06-14  
PRIOR APPLICATION NUMBER: US 09/835,948  
PRIOR FILING DATE: 2001-04-16  
NUMBER OF SEQ ID NOS: 189  
SEQ ID NO 141  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein  
US-10-172-597-141

Query Match  
Best Local Similarity 57.1%; Score 24; DB 14; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.4e+06;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 1 SLEAKLEM 9

RESULT 6  
US-09-865-548A-96  
Sequence 96; Application US/0986548A  
Publication No. US20030096298A1  
GENERAL INFORMATION:  
APPLICANT: Barnea, Elion  
APPLICANT: Beer, Ilan  
APPLICANT: Ziv, Tamara  
APPLICANT: Admon, Aris  
TITLE OF INVENTION: METHOD OF IDENTIFYING PEPTIDES CAPABLE OF BINDING TO MHC MOLECULE  
TITLE OF INVENTION: PEPTIDES IDENTIFIED THEREBY AND THEIR USES  
FILE REFERENCE: 01/22080  
CURRENT APPLICATION NUMBER: US/09/865,548A  
CURRENT FILING DATE: 2001-05-16  
PRIOR APPLICATION NUMBER: US 60/290,958  
PRIOR FILING DATE: 2001-05-16  
NUMBER OF SEQ ID NOS: 204  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 96  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic peptide  
US-09-865-548A-96

Query Match  
Best Local Similarity 54.8%; Score 23; DB 10; Length 9;  
Best Local Similarity 44.4%; Pred. No. 1.4e+06;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 1 SLEAKLEM 9

Db 1 ALTGHEEV 9

## RESULT 7

US-10-008-524A-78  
; Sequence 78, Application US/10008524A  
; Publication No. US20030175682A1  
; GENERAL INFORMATION:  
; APPLICANT: Doobar, John  
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
; FILE REFERENCE: 18396/1074  
; CURRENT APPLICATION NUMBER: US/10/008,524A  
; CURRENT FILING DATE: 2002-08-13  
; PRIOR APPLICATION NUMBER: 09/314,268  
; PRIOR FILING DATE: 1999-05-18  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 78  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-008-524A-78

## Query Match

Best Local Similarity 47.6%; Score 20; DB 14; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LTANKE 8

Db 2 LTAHKT 8

## RESULT 8

US-10-008-524A-79  
; Sequence 79, Application US/10008524A  
; Publication No. US20030175682A1  
; GENERAL INFORMATION:  
; APPLICANT: Doobar, John  
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
; FILE REFERENCE: 18396/1074  
; CURRENT APPLICATION NUMBER: US/10/008,524A  
; CURRENT FILING DATE: 2002-08-13  
; PRIOR APPLICATION NUMBER: 09/314,268  
; PRIOR FILING DATE: 1999-05-18  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 79  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-008-524A-79

## Query Match

Best Local Similarity 47.6%; Score 20; DB 14; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LTANKE 8

Db 2 LTAHKT 8

## RESULT 9

US-10-008-524A-80  
; Sequence 80, Application US/10008524A  
; Publication No. US20030175682A1

## GENERAL INFORMATION:

APPLICANT: Doobar, John  
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
FILE REFERENCE: 18396/1074  
CURRENT APPLICATION NUMBER: US/10/008,524A  
CURRENT FILING DATE: 2002-08-13  
PRIOR APPLICATION NUMBER: 09/314,268  
PRIOR FILING DATE: 1999-05-18  
NUMBER OF SEQ ID NOS: 179  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 80  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-008-524A-80

## Query Match

Best Local Similarity 47.6%; Score 20; DB 14; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LTANKE 8

Db 1 LTAHKT 7

## RESULT 10

US-10-350-719-78  
; Sequence 78, Application US/10350719  
; Publication No. US20030219726A1  
; GENERAL INFORMATION:  
; APPLICANT: Doobar, John  
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
; FILE REFERENCE: 18396/2162  
; CURRENT APPLICATION NUMBER: US/10/350,719  
; CURRENT FILING DATE: 2003-01-24  
; PRIOR APPLICATION NUMBER: PCT/GB01/01176  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: GB0018140.4  
; PRIOR FILING DATE: 2000-07-24  
; NUMBER OF SEQ ID NOS: 180  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 78  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-350-719-78

## Query Match

Best Local Similarity 47.6%; Score 20; DB 14; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LTANKE 8

Db 2 LTAHKT 8

## RESULT 11

US-10-350-719-79  
; Sequence 79, Application US/10350719  
; Publication No. US20030219726A1  
; GENERAL INFORMATION:  
; APPLICANT: Doobar, John  
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
; FILE REFERENCE: 18396/2162

```

; CURRENT APPLICATION NUMBER: US/10/350,719
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: PCT/GB01/01176
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: GB0018140.4
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 79
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-350-719-79
```

```

Query Match      47.6%; Score 20; DB 14; Length 8;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 LTANKE 8
Db      2 LTAHTKD 8
```

```

RESULT 12
US-10-350-719-80
; Sequence 80, Application US/10350719
; Publication No. US20030219726A1
; GENERAL INFORMATION:
; APPLICANT: Decortat, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/2162
; CURRENT APPLICATION NUMBER: US/10/350,719
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: PCT/GB01/01176
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: GB0018140.4
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 80
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-350-719-80
```

```

Query Match      47.6%; Score 20; DB 14; Length 8;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 LTANKE 8
Db      1 LTAHTKD 7
```

```

RESULT 13
US-09-932-165-74
; Sequence 74, Application US/09932165
; Publication No. US20030134784A1
; GENERAL INFORMATION:
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: CHALLITA-EID, PIA M.
; APPLICANT: FARIS, MARY
; APPLICANT: SAFRAN, DOUGLAS
; APPLICANT: AEFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: HUBERT, RENE
```

```

; APPLICANT: GE, WANGMAO
; APPLICANT: JAKOBOVITS, AVA
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED
; TITLE OF INVENTION: 83P2H3 AND CATP2E11 USEFUL IN TREATMENT AND
; FILE REFERENCE: 51158-20014.00
; CURRENT APPLICATION NUMBER: US/09/932,165
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/226,329
; NUMBER OF SEQ ID NOS: 1508
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 74
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-09-932-165-74
```

```

Query Match      47.6%; Score 20; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 ANMK 7
Db      6 ANMK 9
```

```

RESULT 14
US-09-932-165-460
; Sequence 460, Application US/09932165
; Publication No. US20030134784A1
; GENERAL INFORMATION:
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: CHALLITA-EID, PIA M.
; APPLICANT: FARIS, MARY
; APPLICANT: SAFRAN, DOUGLAS
; APPLICANT: AEFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: HUBERT, RENE
; APPLICANT: GE, WANGMAO
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED
; TITLE OF INVENTION: 83P2H3 AND CATP2E11 USEFUL IN TREATMENT AND
; FILE REFERENCE: 51158-20014.00
; CURRENT APPLICATION NUMBER: US/09/932,165
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/226,329
; NUMBER OF SEQ ID NOS: 1508
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 460
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-09-932-165-460
```

```

Query Match      47.6%; Score 20; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 ANMK 7
Db      6 ANMK 9
```

```

RESULT 15
US-09-932-165-664
; Sequence 664, Application US/09932165
```

Publication No. US2003013478A1  
GENERAL INFORMATION:  
APPLICANT: RAITANO, ARTHUR  
APPLICANT: CHALLITA-ETD, PIA M.  
APPLICANT: PARIS, MARY  
APPLICANT: SAFFRAN, DOUGLAS  
APPLICANT: AFEAR, DANIEL  
APPLICANT: LEVIN, ELANA  
APPLICANT: HUBERT, RENE  
APPLICANT: GE, WANGXAO  
APPLICANT: YAKOVITS, AYA  
TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
TITLE OF INVENTION: 83P2H3 AND CATF2E11 USEFUL IN TREATMENT AND  
TITLE OF INVENTION: DETECTION OF CANCER  
FILE REFERENCE: 51158-20014.00  
CURRENT APPLICATION NUMBER: US/09/932,165  
CURRENT FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: 60/226,329  
PRIOR FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 1508  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO: 664  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-664  
  
Query Match 47.6%; Score 20; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 4 ANMK 7  
|||  
Db 6 ANMK 9  
  
Search completed: November 30, 2004, 08:41:16  
Job time : 101 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 08:28:43 ; Search time 20 Seconds  
(without alignments)

29.843 Million cell updates/sec

Title: US-10-008-377A-7

Perfect score: 42

Sequence: 1 SLTNMKEV 9

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 95011

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents, AA:  
1: /cgn2\_6/prodata/1/aa/5A.COMB.pep.\*  
2: /cgn2\_6/prodata/1/aa/5B.COMB.pep.\*  
3: /cgn2\_6/prodata/1/aa/6A.COMB.pep.\*  
4: /cgn2\_6/prodata/1/aa/6B.COMB.pep.\*  
5: /cgn2\_6/prodata/1/aa/6C.COMB.pep.\*  
6: /cgn2\_6/prodata/1/aa/6D.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	57.1	9	US-09-492-543-84	Sequence 84, Appl
2	57.1	9	US-09-492-543-141	Sequence 141, Appl
3	47.6	8	US-08-144-779C-31	Sequence 31, Appl
4	47.6	8	US-09-314-268-78	Sequence 78, Appl
5	47.6	8	US-09-314-268-79	Sequence 79, Appl
6	47.6	8	US-09-314-268-80	Sequence 80, Appl
7	47.6	8	US-09-368-449B-31	Sequence 31, Appl
8	47.6	9	US-08-144-779C-27	Sequence 27, Appl
9	47.6	9	US-09-368-449B-27	Sequence 27, Appl
10	45.2	6	US-08-050-482A-18	Sequence 18, Appl
11	45.2	6	US-08-050-482A-12	Sequence 12, Appl
12	45.2	8	US-08-718-905-8	Sequence 8, Appl
13	45.2	8	US-08-160-604-53	Sequence 53, Appl
14	45.2	8	US-09-550-497-8	Sequence 8, Appl
15	45.2	8	US-09-147-992-8	Sequence 8, Appl
16	45.2	8	US-09-549-839A-8	Sequence 8, Appl
17	45.2	8	US-08-475-955-154	Sequence 154, Appl
18	45.2	8	US-09-780-718-4	Sequence 4, Appl
19	45.2	9	US-09-357-293-4	Sequence 4, Appl
20	45.2	9	US-09-647-378A-48	Sequence 48, Appl
21	45.2	6	US-09-008-308-44	Sequence 44, Appl
22	45.2	6	US-09-256-237-44	Sequence 44, Appl
23	45.2	7	US-08-408-858A-55	Sequence 55, Appl
24	45.2	7	US-09-017-043A-2	Sequence 2, Appl
25	45.2	8	US-08-408-858A-61	Sequence 61, Appl
26	45.2	8	US-08-408-858A-62	Sequence 62, Appl
27	45.2	8	US-09-248-588-66	Sequence 66, Appl

28	18	42.9	8	US-09-206-947-79	Sequence 79, Appl
29	18	42.9	8	US-09-239-043D-583	Sequence 583, Appl
30	18	42.9	8	US-09-239-043D-599	Sequence 599, Appl
31	18	42.9	8	US-09-239-043D-1874	Sequence 1874, Appl
32	18	42.9	8	US-09-239-043D-1910	Sequence 1910, Appl
33	18	42.9	9	US-08-408-858A-68	Sequence 68, Appl
34	18	42.9	9	US-08-408-858A-69	Sequence 69, Appl
35	18	42.9	9	US-09-239-043D-353	Sequence 353, Appl
36	18	42.9	9	US-09-239-043D-600	Sequence 600, Appl
37	18	42.9	9	US-09-239-043D-868	Sequence 868, Appl
38	18	42.9	9	US-09-239-043D-1425	Sequence 1425, Appl
39	18	42.9	9	US-09-239-043D-1911	Sequence 1911, Appl
40	18	42.9	9	US-09-239-043D-1981	Sequence 1981, Appl
41	18	42.9	9	US-09-239-043D-2249	Sequence 2249, Appl
42	17	40.5	5	5217869-93	Patent No. 5217869
43	17	40.5	6	US-08-387-156-14	Sequence 14, Appl
44	17	40.5	6	US-08-742-256-10	Sequence 10, Appl
45	17	40.5	6	US-08-694-865-14	Sequence 14, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-492-543-84
; Sequence 84, Application US/09492543A
; Patent No. 6316213
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-B
; CURRENT APPLICATION NUMBER: US/09/492,543A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: WORD 6.0.1 for Macintosh
; SEQ ID NO 84
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein
US-09-492-543-84

Query Match      57.1%   Score 24, DB 3, Length 9;
Best Local Similarity 55.6%   Pred. No. 3.8e+05;
Matches          5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 SLTNMKEV 9
      |||
      1 SLTNMKEV 9
      |||
      1 SLTNMKEV 9

DB      1 SLTNMKEV 9
      |||
      1 SLTNMKEV 9

RESULT 2
US-09-492-543-141
; Sequence 141, Application US/09492543A
; Patent No. 6316213
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-B
; CURRENT APPLICATION NUMBER: US/09/492,543A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: WORD 6.0.1 for Macintosh
; SEQ ID NO 141
; LENGTH: 9
; TYPE: PRT
```

ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein  
US-09-492-543-141

Query Match 57.1%; Score 24; DB 3; Length 9;  
Best Local Similarity 55.6%; Pred. No. 3.8e+05;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LTNANKEV 9  
DB 1 LTAZKXEM 9

RESULT 3  
US-08-144-779C-31  
Sequence 31, Application US/08144779C  
Patent No. 6068622  
GENERAL INFORMATION:  
APPLICANT: Green, Lawrence R.  
APPLICANT: Sinackevich, Nicolay V.  
APPLICANT: Ivanov, Vadim T.  
APPLICANT: Mikhailova, Inessa I.  
APPLICANT: Vaskovsky, Boris V.  
APPLICANT: Mikhailov, Alexander N.  
APPLICANT: Khavinson, Vladimir K.  
APPLICANT: Morozov, Vyacheslav G.  
TITLE OF INVENTION: Immunomodulating Peptides and Methods  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08144,779C  
FILING DATE: 28-OCT-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/783,517  
FILING DATE: 28-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/816,205  
FILING DATE: 02-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/967,633  
FILING DATE: 28-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Storella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 15542-S-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-144-779C-31

Query Match 47.6%; Score 20; DB 3; Length 8;  
Best Local Similarity 62.5%; Pred. No. 3.8e+05;

Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 LTNANKEV 9  
DB 1 LTAZKXAV 8

RESULT 4  
US-09-314-268-78  
Sequence 78, Application US/09314268  
Patent No. 6346377  
GENERAL INFORMATION:  
APPLICANT: Doobatz, John  
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
FILE REFERENCE: 3789/80902  
CURRENT APPLICATION NUMBER: US/09/314,268  
CURRENT FILING DATE: 1999-03-19  
EARLIER FILING DATE: 1999-05-18  
NUMBER OF SEQ ID NOS: 179  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 78  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: octapeptide antigen  
US-09-314-268-78

Query Match 47.6%; Score 20; DB 3; Length 8;  
Best Local Similarity 57.1%; Pred. No. 3.8e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTNANKE 8  
DB 2 LTAHTKD 8

RESULT 5  
US-09-314-268-79  
Sequence 79, Application US/09314268  
Patent No. 6346377  
GENERAL INFORMATION:  
APPLICANT: Doobatz, John  
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
FILE REFERENCE: 3789/80902  
CURRENT APPLICATION NUMBER: US/09/314,268  
CURRENT FILING DATE: 1999-03-19  
EARLIER FILING DATE: 1999-05-18  
NUMBER OF SEQ ID NOS: 179  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 79  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: octapeptide antigen  
US-09-314-268-79

Query Match 47.6%; Score 20; DB 3; Length 8;  
Best Local Similarity 57.1%; Pred. No. 3.8e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTNANKE 8  
DB 2 LTAHTKD 8



RESULT 6  
US-09-314-268-80  
Sequence 80, Application US/09314268  
Patent No. 6346377  
GENERAL INFORMATION:  
APPLICANT: Doobar, John  
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
FILE REFERENCE: 3789/80902  
CURRENT APPLICATION NUMBER: US/09/314,268  
CURRENT FILING DATE: 1999-03-19  
EARLIER APPLICATION NUMBER: 09/314,268  
EARLIER FILING DATE: 1999-05-18  
NUMBER OF SEQ. ID NOS: 179  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 80  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-314-268-80

Query Match 47.6%; Score 20; DB 3; Length 8;  
Best Local Similarity 57.1%; Pred. No. 3.8e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

CY 2 LTANKE 8  
DB 1 LTAHMD 7

RESULT 7  
US-09-368-449B-31  
Sequence 31, Application US/09368449B  
Patent No. 6346514  
GENERAL INFORMATION:  
APPLICANT: Green, Lawrence R.  
Sinackevich, Nicolay V.  
Ivanov, Vadim T.  
Mikhalyova, Inessa I.  
Vaskovsky, Boris V.  
Mikhaltsov, Alexander N.  
Khavinson, Vladimir K.  
Morozov, Vyacheslav G.  
TITLE OF INVENTION: Immunomodulating Peptides and Methods  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/368,449B  
FILING DATE: 04-Aug-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/144,779  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/816,205  
FILING DATE: 02-JAN-1992  
APPLICATION NUMBER: US 07/967,633  
FILING DATE: 28-OCT-1992  
ATTORNEY/AGENT INFORMATION:

NAME: Storella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 15542-5-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 31:  
US-09-368-449B-31

Query Match 47.6%; Score 20; DB 3; Length 8;  
Best Local Similarity 62.5%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

CY 2 LTANKEV 9  
DB 1 LTAZKAV 8

RESULT 8  
US-08-144-779C-27  
Sequence 27, Application US/08144779C  
Patent No. 606622  
GENERAL INFORMATION:  
APPLICANT: Green, Lawrence R.  
Sinackevich, Nicolay V.  
Ivanov, Vadim T.  
Mikhalyova, Inessa I.  
Vaskovsky, Boris V.  
Mikhaltsov, Alexander N.  
Khavinson, Vladimir K.  
Morozov, Vyacheslav G.  
TITLE OF INVENTION: Immunomodulating Peptides and Methods  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/144,779C  
FILING DATE: 28-OCT-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/783,517  
FILING DATE: 28-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/816,205  
FILING DATE: 02-JAN-1992  
APPLICATION NUMBER: US 07/967,633  
FILING DATE: 28-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Storella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 15542-5-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 27;  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-144-779C-27

Query Match 47.6%; Score 20; DB 3; Length 9;  
Best Local Similarity 62.5%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LTMKEV 9  
DB 2 LTAZKAV 9

RESULT 9  
US-09-368-449B-27  
Sequence 27, Application US/09368449B  
Patent No. 6346514  
GENERAL INFORMATION:  
APPLICANT: Green, Lawrence R.  
Sinackevich, Nicolay V.  
Ivanov, Vadim T.  
Mikhalyova, Inessa I.  
Vaskovsky, Boris V.  
Mikhailtsov, Alexander N.  
Khavinson, Vladimir K.  
Morozov, Vyacheslav G.  
TITLE OF INVENTION: Immunomodulating Peptides and Methods of Use  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/368,449B  
FILING DATE: 04-Aug-1999  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/144,779  
FILING DATE: <unknown>  
APPLICATION NUMBER: US 07/816,205  
FILING DATE: 02-JAN-1992  
APPLICATION NUMBER: US 07/967,633  
FILING DATE: 28-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Storella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 15542-5-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: <unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-09-368-449B-27

Query Match 47.6%; Score 20; DB 3; Length 9;  
Best Local Similarity 62.5%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LTMKEV 9  
DB 2 LTAZKAV 9

RESULT 10  
US-08-050-482A-18  
Sequence 18, Application US/08050482A  
Patent No. 6312939  
GENERAL INFORMATION:  
APPLICANT: ROBERTS, Joseph  
MACALLISTER, Thomas W.  
SETHURAMAN, Natarajan  
FREEMAN, Abbie G.  
TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND ITS USE IN ANTIVIRAL AND ANTICANCER THERAPY  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/050,482A  
FILING DATE: 25-Apr-1995  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US92/10421  
FILING DATE: 04-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 023032/0106  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5399  
TELEFAX: (202) 672-5399  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-08-050-482A-18

Query Match 45.2%; Score 19; DB 3; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 1 MKEV 4

RESULT 11  
US-08-050-482A-22  
Sequence 22, Application US/08050482A  
Patent No. 6312939  
GENERAL INFORMATION:  
APPLICANT: ROBERTS, Joseph  
MACALLISTER, Thomas W.

SETHURAMAN, Natarajan  
FREEMAN, Abbie G.  
TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND ITS USE IN ANTIVIRAL AND ANTICANCER THERAPY  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: POLEY & LARDNER  
STREET: 3000 K Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/050,482A  
FILING DATE: 25-Apr-1995  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US92/10421  
FILING DATE: 04-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A.  
REGISTRATION NUMBER: 23,768  
REFERENCE/DOCKET NUMBER: 023032/0106  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-08-050-482A-22  
Query Match 45.2%; Score 19; DB 3; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 6 MKEY 9  
DB 1 MKEY 4  
RESULT 12  
US-08-718-905-8  
Sequence 8, Application US/08718905  
Patent No. 6063755  
GENERAL INFORMATION:  
APPLICANT: Donovan, William P.  
APPLICANT: Donovan, Judith C.  
APPLICANT: Slaney, Annette C.  
TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET33 AND CRYET34  
TITLE OF INVENTION: COMPOSITIONS AND USES THEREFOR  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/718,905  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: MOBT:003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-718-905-8  
Query Match 45.2%; Score 19; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 6 MKEY 9  
DB 2 MKEY 5  
RESULT 13  
US-08-160-604-53  
Sequence 53, Application US/08160604  
Patent No. 6232522  
GENERAL INFORMATION:  
APPLICANT: Harley, John  
APPLICANT: James, Judith A.  
APPLICANT: Scottfield, R. H.  
TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY  
NUMBER OF SEQUENCES: 127  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30308-4530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/160,604  
FILING DATE: 30-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/867,819  
FILING DATE: 13-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/648,205  
FILING DATE: 31-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,947  
FILING DATE: 31-JAN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRE114CIP(3)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)-815-6558  
TELEFAX: (404)-815-6555  
INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-160-604-53

Query Match 45.2% Score 19; DB 3; Length 8;  
Best Local Similarity 50.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LTKMKEV 9  
DB 1 LTKMKEV 8

RESULT 14  
US-09-550-497-8

Sequence 8, Application US/09550497  
Patent No. 6248536  
GENERAL INFORMATION:

APPLICANT: Donovan, William P.  
Staney, Judith C.  
TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET33 AND CRYET34

COMPOSITIONS AND USES THEREFOR  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/550,497  
FILING DATE: 14-Apr-2000  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/718,905  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: MOBT.003

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-550-497-8

Query Match 45.2% Score 19; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 2 MKEV 5

RESULT 15

US-09-147-992-8

Sequence 8, Application US/09147992  
Patent No. 6326351  
GENERAL INFORMATION:

APPLICANT: BACILLUS THURINGIENSIS CRYET33 AND CRYET34  
TITLE OF INVENTION: COMPOSITIONS AND USES THEREFOR  
NUMBER OF SEQUENCES: 10  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/147,992  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/718,905  
FILING DATE: 24-SEP-1996  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

Query Match 45.2% Score 19; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 2 MKEV 5

US-09-147-992-8

Sequence 8, Application US/09147992  
Patent No. 6326351  
GENERAL INFORMATION:

APPLICANT: BACILLUS THURINGIENSIS CRYET33 AND CRYET34  
TITLE OF INVENTION: COMPOSITIONS AND USES THEREFOR  
NUMBER OF SEQUENCES: 10  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/147,992  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/718,905  
FILING DATE: 24-SEP-1996  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-09-147-992-8

Query Match 45.2% Score 19; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 2 MKEV 5

Search completed: November 30, 2004, 08:35:48  
Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: November 30, 2004, 07:41:32 / Search time 11.6667 Seconds  
(without alignments)  
82.472 Million cell updates/sec

Title: US-10-008-377A-6  
Percent score: 52  
Sequence: 1 ALPIEXGPLV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 263416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: PIR 79:.\*  
1: p1r1:.\*  
2: p1r2:.\*  
3: p1r3:.\*  
4: p1r4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	38.5	7	2 A44428	platelet aggregati
2	18	34.6	9	2 S63491	dissimulatory sulf
3	18	34.6	9	2 S70332	endospem protein,
4	17	32.7	8	2 PT0311	Ig heavy chain CRD
5	17	32.7	8	2 E47393	neuropeptide calla
6	17	32.7	8	2 A28719	thymic humoral fac
7	17	32.7	9	2 G56978	collagen alpha 1(I
8	17	32.7	10	2 PC2171	triacylglycerol 11
9	16	30.8	5	2 UT0520	Ig kappa chain V-I
10	16	30.8	7	2 I48086	DNA topoisomerase
11	16	30.8	8	2 S15422	adipokinetic hormo
12	16	30.8	8	2 A58641	adipokinetic hormo
13	16	30.8	8	2 PT0368	Ig gamma chain C r
14	16	30.8	8	2 S21653	neuropeptide - flo
15	16	30.8	10	2 S65388	Cytochrome-c oxida
16	16	30.8	10	2 C39398	Fc mu (IgM) recept
17	16	30.8	10	2 PH1633	Ig H chain V-D-J r
18	15.5	29.8	10	2 C39111	Ig heavy chain C r
19	15	28.8	5	2 UN0860	peptidyl-dipeptida
20	15	28.8	6	2 A61049	halo-toxin - Pseud
21	15	28.8	7	2 S66442	glutathione S-tran
22	15	28.8	7	2 B48394	major fat-globule
23	15	28.8	8	2 D47393	neuropeptide calla
24	15	28.8	9	2 PT0285	Ig heavy chain CRD
25	15	28.8	9	2 S36850	Ig heavy chain V r
26	15	28.8	9	2 PT0080	60k Ca binding pro
27	15	28.8	9	2 PD0443	3-oxoacid CoA-tran
28	15	28.8	10	1 XASNP2	angiotensin-conver
29	15	28.8	10	1 GMROL2	leucosulfakinin-II

30	15	28.8	10	2 B55482	platelet activatin
31	15	28.8	10	2 B60656	leucosulfakinin II
32	15	28.8	10	2 S59030	lyso-bradykinin -
33	14.5	27.9	9	2 G58502	kidney and bladder
34	14	26.9	5	2 S60274	major protein anti
35	14	26.9	5	2 PQ0689	photosystem I 10.4
36	14	26.9	7	2 I48105	dihydrofolate redu
37	14	26.9	7	2 ECMUCR	catch-relaxing pep
38	14	26.9	8	2 C39650	neural cell adhesi
39	14	26.9	9	2 S66698	quinoline-2-oxidor
40	14	26.9	9	2 B49712	ATP-binding protei
41	14	26.9	9	2 S26508	collagen alpha 2(V
42	14	26.9	9	2 B60246	ornitho-kinin - ch
43	14	26.9	10	1 RHLMGS	gonadoliberin - se
44	14	26.9	10	2 C26997	unspecific monooxy
45	14	26.9	10	2 P50451	24k protein 4302 -

## ALIGNMENTS

RESULT 1  
A44428  
platelet aggregation-associated protein - Streptococcus sanguis (fragment)  
C:Species: Streptococcus sanguis  
C:Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 11-Nov-1994  
C:Accession: A44428  
R:Ericsson, P.R.; Herzberg, M.C.  
J. Biol. Chem. 268, 1646-1649, 1993  
A:Title: The Streptococcus sanguis platelet aggregation-associated protein. Identificat:  
A:Reference number: A44428; MUID:91131902; PMID:8420939  
A:Accession: A44428  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <ERI>  
C:Keywords: cell wall

Query Match  
Best Local Similarity 38.5%; Score 20; DB 2; Length 7;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 2  
S63491  
dissimulatory sulfite reductase beta chain, membrane-bound - Desulfovibrio desulfuricans  
C:Species: Desulfovibrio desulfuricans  
C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
C:Accession: S63491  
R:Stender, J.; Arendsen, A.F.; Hagen, W.R.; Kroneck, P.M.H.  
Eur. J. Biochem. 233, 873-879, 1995  
A:Title: Molecular properties of the dissimulatory sulfite reductase from Desulfovibrio  
A:Reference number: S63491; MUID:9608152; PMID:8521853  
A:Accession: S63491  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <STR>

Query Match  
Best Local Similarity 34.6%; Score 18; DB 2; Length 9;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 P1EXGP 8  
DB 4 P1GVNP 9

RESULT 3  
S70332  
endospem protein, 10K - rye (fragment)

C:Species: Secale cereale (rye)  
 C>Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #next\_change 17-Apr-1998  
 C:Accession: S70332  
 R:Roche, A.; Calero, M.; Soriano, F.; Mendez, E.  
 Biochem. Biophys. Acta 1295, 13-22, 1996  
 A:Title: Identification of major rye secalins as coeliac immunoreactive proteins.  
 A:Reference number: S70327; MUID:56283789; PMID:8679669  
 A:Accession: S70332  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-9 <RCC>

Query Match 34.6%; Score 18; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 PLEV 6  
 ||||  
 DB 3 PLEV 6

RESULT 4  
 P70311  
 Ig heavy chain CDR3 region (clone 6-100) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #next\_change 16-Aug-1996  
 C:Accession: P70311  
 R:Yamada, M.; Masserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
 A:Reference number: P70222; MUID:91108337; PMID:1899102  
 A:Accession: P70311  
 A:Molecule type: DNA  
 A:Residues: 1-8 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterodimer; immunoglobulin

Query Match 32.7%; Score 17; DB 2; Length 8;  
 Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 PLEV 7  
 ||||  
 DB 1 PSAYG 5

RESULT 5  
 E47393  
 neuropeptide callistostatin 5 - blowbottle fly (Calliphora vomitoria)  
 C:Species: Calliphora vomitoria  
 C>Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #next\_change 09-Jul-2004  
 C:Accession: E47393  
 R:Duvé, H.; Johnson, A.H.; Scott, A.G.; Yu, C.G.; Yagci, K.J.; Tobe, S.S.; Thorpe, A.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 2456-2460, 1993  
 A:Title: Callistostatin: neuropeptides from the blowfly Calliphora vomitoria with sequen  
 A:Reference number: A47393; MUID:93211980; PMID:8460157  
 A:Accession: E47393  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-8 <DUV>  
 A:Cross-references: UNIPROT:P41841  
 A:Experimental source: whole flies  
 A:Note: sequence extracted from NCI backbone (NCBIP:128482)

Query Match 32.7%; Score 17; DB 2; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 PLEV 7  
 ||||  
 DB 3 PYDFG 7

RESULT 6  
 A28719  
 thymic humoral factor gamma-2 - bovine (fragment)  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #next\_change 09-Jul-2004  
 C:Accession: A28719  
 R:Birstein, Y.; Buchner, V.; Pecht, M.; Trainin, N.  
 Biochemistry 27, 4066-4071, 1988  
 A:Title: Thymic humoral factor gamma-2: purification and amino acid sequence of an immu  
 A:Reference number: A28719; MUID:88326920; PMID:3261994  
 A:Accession: A28719  
 A:Molecule type: protein  
 A:Residues: 1-8 <BUR>  
 A:Cross-references: UNIPROT:O18787

Query Match 32.7%; Score 17; DB 2; Length 8;  
 Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 IEXGP 8  
 ||||  
 DB 1 LEDGP 5

RESULT 7  
 G56978  
 collagen alpha 1(II) chain - bovine (fragment)  
 N:Alternate names: collagen alpha 3(XI) chain  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #next\_change 03-Oct-1995  
 C:Accession: G56978  
 R:Wu, J.C.; Eyre, D.R.  
 J. Biol. Chem. 270, 18865-18870, 1995  
 A:Title: Structural analysis of cross-linking domains in cartilage type XI collagen. In  
 A:Reference number: A56978; MUID:95370194; PMID:7642541  
 A:Accession: G56978  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-9 <WU>  
 A:Note: the residue designated 'X' is modified lysine in collagen 3(XI) some cross-link

Query Match 32.7%; Score 17; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 EYGP 8  
 ||||  
 DB 4 EXGP 7

RESULT 8  
 PC2171  
 triacylglycerol lipase (EC 3.1.1.3) I A chain - Rhizopus niveus (strain IFO 4759) (fragm  
 C:Species: Rhizopus niveus  
 C>Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #next\_change 09-Jul-2004  
 C:Accession: PC2171  
 R:Kohno, M.; Kugimiya, W.; Hashimoto, Y.; Morita, Y.  
 Biosci. Biotechnol. Biochem. 58, 1007-1012, 1994  
 A:Title: Purification, characterization, and crystallization of two types of lipase from  
 A:Reference number: PC2171; MUID:94319059; PMID:7765029  
 A:Accession: PC2171  
 A:Molecule type: protein  
 A:Residues: 1-10 <KOH>  
 A:Cross-references: UNIPROT:O74166  
 C:Comment: This enzyme catalyzes the hydrolysis of the ester bonds.  
 C:Keywords: carboxylic ester hydrolase

Query Match 32.7%; Score 17; DB 2; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 3.4e+03;  
 Matches 5; Conservative 1; Mismatches 0; Indels 4; Gaps 1;

OY 1 ALPIEYGPV 10  
 |||||

Db 4 ALP----PLI 9

## RESULT 9

IG kappa chain V-II region (SD1) - human (fragment)  
 C/Species: Homo sapiens (man)  
 C/Date: 23-Oct-1992 #sequence\_revision 23-Oct-1992 #text\_change 16-Aug-1996  
 C/Accession: J0520  
 R/Anker: R.; Conley, M.E.; Pollok, B.A.  
 J. Exp. Med. 169, 2109-2119, 1989  
 A/Title: Clonal diversity in the B cell repertoire of patients with X-linked agammaglobulinemia  
 A/Reference number: J0511, MUID:89279157, PMID:2786547  
 A/Accession: J0520  
 A/Molecule type: mRNA  
 A/Residues: 1-5 <ANK>  
 A/Note: the sequence shown here is one of eight productive V-D-J mu chain rearrangements  
 A/Note: a stop codon terminates the sequence in the V region  
 C/Keywords: heterotetramer; immunoglobulin  
 F:1-5/Domain: V kappa region <VRB>

Query Match 30.8%; Score 16; DB 2; Length 5;  
 Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 EYGP 8  
 :  
 :  
 :  
 Db 2 EGGP 5

## RESULT 10

DNA topoisomerase II alpha - Chinese hamster (fragment)  
 C/Species: Cricetus griseus (Chinese hamster)  
 C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
 C/Accession: I48086  
 R/NG: S.W.; Eder, J.P.; Schnipper, L.E.; Chan, V.T.W.  
 J. Biol. Chem. 270, 25850-25858, 1995  
 A/Title: Molecular cloning and characterization of the promoter for the Chinese hamster  
 A/Reference number: I48086, MUID:96029664, PMID:7592770  
 A/Accession: I48086  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-7 <RES>  
 A/Cross-references: EMBL:U34196; MUD:91041231; PIDN:AC52315.1; PID:91041232

Query Match 30.8%; Score 16; DB 2; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEYGP 9  
 :  
 :  
 :  
 Db 1 MELSP 6

## RESULT 11

adipokinetic hormone - cockchafer  
 C/Species: Melolontha melolontha (cockchafer)  
 C/Date: 19-Mar-1997 #sequence\_revision 31-Oct-1997 #text\_change 09-Jul-2004  
 C/Accession: S15422  
 R/Gade, G.  
 Biochem. J. 275, 671-677, 1991  
 A/Title: A unique charged tyrosine-containing member of the adipokinetic hormone/red-pigment concentrating hormone family  
 A/Reference number: S15422, MUID:91248100, PMID:2039445  
 A/Accession: S15422  
 A/Molecule type: protein  
 A/Residues: 1-8 <BIO>  
 A/Cross-references: UNIPROT:P25423  
 C/Superfamily: adipokinetic hormone  
 C/Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid  
 F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
 F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 30.8%; Score 16; DB 2; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEYGP 8  
 :  
 :  
 :  
 Db 2 LNYGP 6

## RESULT 12

adipokinetic hormone - dor beetle  
 C/Species: Geotrupes stercorarius (dor beetle)  
 C/Date: 28-Oct-1997 #sequence\_revision 31-Oct-1997 #text\_change 09-Jul-2004  
 C/Accession: A58641  
 R/Gade, G.  
 Biochem. J. 275, 671-677, 1991  
 A/Title: A unique charged tyrosine-containing member of the adipokinetic hormone/red-pigment concentrating hormone family  
 A/Reference number: S15422, MUID:91248100, PMID:2039445  
 A/Accession: A58641  
 A/Molecule type: protein  
 A/Residues: 1-8 <BIO>  
 A/Cross-references: UNIPROT:P25423  
 C/Superfamily: adipokinetic hormone  
 C/Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid  
 F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
 F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 30.8%; Score 16; DB 2; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEYGP 8  
 :  
 :  
 :  
 Db 2 LNYGP 6

## RESULT 13

IG gamma chain C region (gamma-1) - human (fragment)  
 C/Species: Homo sapiens (man)  
 C/Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Aug-1996  
 C/Accession: P0368  
 R/Mullis, M.; Fougereau, M.; Guillelmi, P.; Schiff, C.  
 Mol. Immunol. 28, 753-761, 1991  
 A/Title: Early occurrence of immunoglobulin isotype switching in human fetal liver  
 A/Reference number: P0368, MUID:9132348, PMID:1906381  
 A/Accession: P0368  
 A/Molecule type: mRNA  
 A/Residues: 1-8 <MIL>  
 A/Experimental source: fetal liver  
 C/Keywords: immunoglobulin

Query Match 30.8%; Score 16; DB 2; Length 8;  
 Best Local Similarity 33.3%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEYGP 9  
 :  
 :  
 :  
 Db 1 LHOGP 6

## RESULT 14

neuropeptide - flower beetle (Pachnoda marginata)  
 C/Species: Pachnoda marginata  
 C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
 C/Accession: S21663  
 R/Gade, G.; Lopata, A.; Kellner, R.; Rinehart, K.L.  
 Biol. Chem. Hoppe-Seyler 373, 133-142, 1992  
 A/Title: Primary structures of neuropeptides isolated from the corpora cardiaca of various insects

A:Reference number: S21663; MUID:92265187; PMID:1586453  
 A:Accession: S21663  
 A:Status: Preliminary  
 A:Molecule type: protein  
 A:Residues: 18 <GRF>  
 A:Cross-references: UNIPROT:P25423

Query Match 30.8%; Score 16; DB 2; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEYGP 8  
 : | | |  
 DB 2 LNYSP 6

RESULT 15  
 S65388  
 cytochrome-c oxidase (EC 1.9.3.1) chain VII c, hepatic - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C:Accession: S65388; S65389  
 R:Schlegger, H.; Noack, H.; Halangk, W.; Brandt, U.; von Jagow, G.  
 Eur. J. Biochem. 230, 235-241, 1995  
 A:Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-term  
 A:Reference number: S65372; MUID:95324529; PMID:7601105  
 A:Accession: S65388  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-10 <SCH>  
 A:Cross-references: UNIPROT:P80432  
 A:Accession: S65389  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-10 <SC2>  
 C:Superfamily: cytochrome-c oxidase chain VIIc  
 C:Keywords: oxidoreductase

Query Match 30.8%; Score 16; DB 2; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 5.4e+03;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 EYGP 8  
 : | | |  
 DB 4 EGGP 7

Search completed: November 30, 2004, 08:14:27  
 Job time : 12.6667 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:40:12 / Search time 78.333 Seconds  
(without alignments)  
73.452 Million cell updates/sec

Title: US-10-008-377a-6  
Perfect score: 52  
Sequence: 1 ALPIEYGPV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 2971

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: UniProt\_02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23.5	45.2	10	2 P83154	P83154 anabaena sp
2	20	38.5	9	2 Q7RSP2	Q7RSP2 plasmodium
3	19	36.5	8	2 Q9R7H9	Q9R7H9 haemophilus
4	18	36.5	8	1 ALI6 CARMA	P81819 carcinoma ma
5	16	34.6	8	2 Q16468	Q16468 homo sapien
6	16	34.6	8	2 Q83332	Q83332 murine hepa
7	17	32.7	7	1 MNPI_LBPDE	P43984 leptinotars
8	17	32.7	8	1 ALI4_CYPPO	P82155 cydia pomon
9	17	32.7	8	1 ALI5_CALVO	P41841 calliphora
10	17	32.7	8	2 Q7KYV5	Q7KYV5 homo sapien
11	17	32.7	8	2 Q9X3X1	Q9X3X1 prochloroco
12	17	32.7	9	2 Q9TWV0	Q9TWV0 anthoplaura
13	17	32.7	10	2 Q6JVP0	Q6JVP0 woliastonia
14	17	32.7	10	2 Q6JVP6	Q6JVP6 scopapnus e
15	17	32.7	10	2 Q6JVP6	Q6JVP6 elaphandra
16	17	32.7	10	2 Q6JVP8	Q6JVP8 dimerosstem
17	17	32.7	10	2 Q6JVP0	Q6JVP0 dimerosstem
18	17	32.7	10	2 Q6JVP3	Q6JVP3 baillimora x
19	17	32.7	10	2 Q6JVP5	Q6JVP5 angelphytum
20	17	32.7	10	2 Q6JVP7	Q6JVP7 angelphytum
21	17	32.7	10	2 Q96041	Q96041 cenothera p
22	17	32.7	10	2 Q8SH93	Q8SH93 brookesia b
23	17	32.7	10	2 Q9R7J8	Q9R7J8 helicobacte
24	17	32.7	10	2 Q9GVJ5	Q9GVJ5 rattus sp.
25	17	32.7	10	2 Q9GVJ6	Q9GVJ6 rattus sp.
26	17	32.7	10	2 AARI4752	AARI4752 angelphyc
27	17	32.7	10	2 AARI4754	AARI4754 angelphyc
28	17	32.7	10	2 AARI4756	AARI4756 baillimora
29	17	32.7	10	2 AARI4760	AARI4760 dimerosste
30	17	32.7	10	2 AARI4762	AARI4762 dimerosste
31	17	32.7	10	2 AARI4764	AARI4764 elaphandr

32	17	32.7	10	2 AARI4768	AARI4768 ctocarpus
33	17	32.7	10	2 AARI4772	AARI4772 woliaston
34	16	30.8	5	1 FARP_CHICK	P83308 gallus gall
35	16	30.8	6	1 SAPP_SEPOF	P83569 sepia offic
36	16	30.8	8	1 AKH_MELMU	P25423 melolontha
37	16	30.8	8	1 ALI7 CARMA	P81809 carcinoma ma
38	16	30.8	8	2 Q7OKG9	Q7OKG9 sus scrofa
39	16	30.8	8	2 P82598	P82598 rattus norv
40	16	30.8	8	2 CAE06499	CAE06499 sus scrof
41	16	30.8	8	2 CAE06500	CAE06500 sus scrof
42	16	30.8	8	2 CAE06501	CAE06501 sus scrof
43	16	30.8	8	2 CAE06502	CAE06502 sus scrof
44	16	30.8	9	2 Q9UE09	Q9UE09 homo sapien
45	16	30.8	9	2 P82568	P82568 streptococc

## ALIGNMENTS

RESULT 1					
ID	P83154	PRELIMINARY;	PRT;	10 AA.	
AC	P83154				
DT	01-MAR-2002 (TREMUREL, 20, Created)				
DT	01-MAR-2002 (TREMUREL, 20, Last sequence update)				
DE	01-OCT-2003 (TREMUREL, 25, Last annotation update)				
DE	Phycobiosome rod-core linker polypeptide cpoc3 (Fragment).				
OS	Anabaena sp. (strain 131)				
OC	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.				
OX	NCBI_TaxID=29412;				
RN	(1)				
RP	SEQUENCE.				
RA	Apte S.K., Uhlemann E., Schmid R., Alenddorf K.				
RL	Submitted (OCT-2001) to Swiss-Prot.				
CC	-1- FUNCTION: Rod-core linker protein required for attachment of				
CC	phycocyanin to allophycocyanin in cores of phycobiosomes.				
CC	-1- FUNCTION: Linker polypeptides determine the state of aggregation				
CC	and the location of the disk-shaped phycoliprotein units within				
CC	the phycobiosome and modulate their spectroscopic properties in				
CC	order to mediate a directed and optimal energy transfer.				
CC	-1- SUBUNIT: The phycobiosome is a hemidiscoidal structure that is				
CC	composed of two distinct substructures: a core complex (that				
CC	contains phycobilliproteins) and a number of rods radiating from				
CC	the core.				
CC	-1- SUBCELLULAR LOCATION: Thylakoid membrane.				
CC	-1- SIMILARITY: TO OTHER PHYCOBILISOME LINKER PROTEINS.				
DR	GO; GO:0016020; C:membrane; IEA.				
DR	GO; GO:0030089; C:phycobiosome; IEA.				
DR	GO; GO:0009579; C:thylakoid; IEA.				
DR	GO; GO:0015979; P:photosynthesis; IEA.				
KW	Membrane; Photosynthesis; Phycobiosome; Thylakoid.				
FT	NON TER				
FT	10				
SQ	SEQUENCE 10 AA; 1144 MW; 2F9B662B5B172737 CRC64;				
Query Match 45.2%; Score 23.5; DB 2; Length 10;					
Best Local Similarity 66.7%; Pred. No. 1.6e+03;					
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;					
QY	1 ALP-IEYGP 8				
DB	1 ALPIEYXP 9				
RESULT 2					
ID	Q7RSP2	PRELIMINARY;	PRT;	9 AA.	
AC	Q7RSP2				
DT	01-MAR-2004 (TREMUREL, 26, Created)				
DT	01-MAR-2004 (TREMUREL, 26, Last sequence update)				
DT	01-MAR-2004 (TREMUREL, 26, Last annotation update)				
DE	Hypothetical protein.				
GN	Name=PY00313;				
OS	Plasmodium yoelii yoelii.				

```

DR NCBI_TaxID=6759;
CC Eumalacostraca, Eucarida, Decapoda; Pleocyemata; Brachyura;
OC Amphichela, Portunoidae; Portundidae; Carcinus.
RN [1]
RP SEQUENCE.
RA TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=946195;
RA Dave H., Johnson A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RL allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur.-J. Biochem. 250:727-734(1997).
CC CC -1- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC CC -1- SIMILARITY: Belongs to the allatostatin family.
RW Annotation: Direct protein sequencing; Multigene family; Neuropeptide.
FT MOD RBS 8
SQ SEQUENCE 8 AA; 813 MW; 7C286B45AB476878 CRC64;

Query Match          36.5%; Score 19; DB 1; Length 8;
Best Local Similarity 60.0%; Pred. No. 1.8e+06;
Matches      3; Conservative      0; Mismatches     2; Indels      0; Gaps      0;

QY      3 PIRYG 7
DB      3 PYSYG 7

RESULT 5
ID O16468 PRELIMINARY; PRT; 8 AA.
AC O16468;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE H.sapiens DNA for cosmid ccl3-1134 PCR primer 1 (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA James L.A., Ogilvie D.J., Yamakawa K., Nakamura Y., Strirling C.J.,
RA Hand R.;
RT "Walking, cloning and mapping with YACs in 3q27. Localisation of 5
RT ESTs including 3 members of the Cystatin gene family and
RT identification of CpG islands.";
RL Genomics 32:425-430(1996).
DL EMBL; X88976; CAA61407.1; -.
FT NON TER 8
SQ SEQUENCE 8 AA; 925 MW; FDS411A7376871E6 CRC64;

Query Match          36.5%; Score 19; DB 2; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.8e+06;
Matches      3; Conservative      1; Mismatches     0; Indels      0; Gaps      0;

QY      6 YGFPL 9
DB      2 HGPL 5

RESULT 6
ID O83332 PRELIMINARY; PRT; 8 AA.
AC O83332;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

```

DE HE (Fragment).  
 OS Murine hepatitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 CC Coronaviridae; Coronavirinae; Group 2 species.  
 CX NCBI\_TaxID=1118;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JHM;  
 RX MEDLINE=95191005; PubMed=7884877;  
 RA Kim K.H., Makino S.;  
 RT "Two murine coronavirus genes suffice for viral RNA synthesis.";  
 RL J. Virol. 69:2313-2321(1995).  
 DR EMBL: U19933; AAA69002.1; -;  
 FT NON\_TER 1 1  
 FT MOD\_RES 8 8  
 FT SEQUENCE 8 AA; 907 MW; 922735B1735A2CD CRC64;

Query Match 34.6%; Score 18; DB 2; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Cy 1 ALPIRYG 6  
 Db 1 AYSLEY 6

RESULT 7  
 NMPL LEPEDE STANDARD; PRT; 7 AA.  
 ID NMPL LEPEDE  
 AC P42984;  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Myotropic neuropetide 1 (led-MNP-I)  
 OS Lepidoptera: decaimnata (Colorado potato beetle).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neuroptera; Megaloptera; Coleoptera; Polyphaga; Cuculliformia;  
 CC Phytomyzidae; Chrysomelidae; Chrysomelinae;  
 CC Chrysomelidae; Leptinotarsa.  
 CC NCBI\_TaxID=7539;  
 RN [1]  
 RP SEQUENCE AND SYNTHESIS.  
 RC TISSUE=Head;  
 RA Spittels K., Vankeerberghen A., Schoofs L., Torckens S.,  
 RA Grauwels L., van Leuven F., de Loof A.;  
 RT "Identification, characterization, and immunological localization of a  
 RT novel myotropic neuropetide in the Colorado potato beetle,  
 RT Lepidoptera decaimnata.";  
 RL Peptides 16:365-374(1995).  
 CC -1- FUNCTION: Myotropic peptide. Stimulates the contractions of the  
 CC oviduct.  
 CC KW Annotation: Direct protein sequencing; Neuropeptide.  
 FT MOD\_RES 7 7  
 FT SEQUENCE 7 AA; 705 MW; 6DD373768745B5DB0 CRC64;

Query Match 32.7%; Score 17; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 7 GPL 9  
 Db 4 GPL 6

RESULT 8  
 ALI4 CYDPO STANDARD; PRT; 8 AA.  
 ID ALI4 CYDPO  
 AC P82155;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Cydactactin 4.

OS Cydia pomonella (Codling moth).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neuroptera; Megaloptera; Lepidoptera; Glossata; Ditrysia;  
 CC Tortricidae; Tortricidae; Olethreutinae; Cydia.  
 CX NCBI\_TaxID=82600;  
 RN [1]  
 RP SEQUENCE  
 RC TISSUE=Larva;  
 RX MEDLINE=98054539; PubMed=9392829;  
 RA Duvé H., Johnson A.H., Maestri U.-L., Scott A.G., Minstamley D.,  
 RA Davey M., East P.D., Thorpe A.;  
 RT "Lepidopteran peptides of the allatostatin superfamily.";  
 RL Peptides 18:1301-1309(1997).  
 CC -1- SIMILARITY: Belongs to the allatostatin family.  
 CC Annotation: Direct protein sequencing; Neuropeptide.  
 FT MOD\_RES 8 8  
 FT SEQUENCE 8 AA; 910 MW; 92879D5A847740D CRC64;

Query Match 32.7%; Score 17; DB 1; Length 8;  
 Best Local Similarity 42.9%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Cy 1 ALPIRYG 7  
 Db 1 ARPYRFG 7

RESULT 9  
 ALIS CALVO STANDARD; PRT; 8 AA.  
 ID ALIS CALVO  
 AC P41841;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Callaostatin 5 (Met-callaostatin 1) ([Hyp3]Met-callaostatin).  
 OS Calliphora vomitoria (Blue blowfly).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neuroptera; Megaloptera; Diptera; Brachycera; Muscomorpha; Oestroidea;  
 CC Calliphoridae; Calliphora.  
 CC NCBI\_TaxID=27454;  
 RN [1]  
 RP SEQUENCE  
 RC TISSUE=Head;  
 RA Duvé H., Johnson A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S.,  
 RA Thorpe A.;  
 RT "Callaostatin: neuropeptides from the blowfly Calliphora vomitoria  
 RT with sequence homology to cockroach allatostatin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).  
 RN [2]  
 RP CHARACTERIZATION, AND HYDROXYLATION.  
 RC TISSUE=Head;  
 RX MEDLINE=9434269; PubMed=8063725;  
 RA Duvé H., Johnson A.H., Scott A.G., East P., Thorpe A.;  
 RT "[Hyp3]Met-callaostatin. Identification and biological properties of  
 RT a novel neuropeptide from the blowfly Calliphora vomitoria.";  
 RL J. Biol. Chem. 269:21059-21066(1994).  
 CC -1- FUNCTION: May act as a neurotransmitter or neuromodulator and play  
 CC a role in the integration of information within the brain. May be  
 CC involved in the control of visceral muscles due to its ability to  
 CC behave as potent inhibitors of peristaltic movements. May also  
 CC fulfill a neurohormonal role on muscles of the gut and heart.  
 CC -1- TISSUE SPECIFICITY: Neurons within brain and abdominal ganglion.  
 CC -1- SIMILARITY: Belongs to the allatostatin family.  
 CC PIR: E47393; E47393.  
 DR PIR: E47393; E47393.  
 KW Annotation: Direct protein sequencing; Hydroxylation; Neuropeptide.  
 FT MOD\_RES 3 3  
 FT MOD\_RES 8 8  
 FT SEQUENCE 8 AA; 883 MW; 7D9679CAB5477768 CRC64;

Query Match 32.7%; Score 17; DB 1; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 PLEVG 7  
DB 3 PYDFG 7

## RESULT 10

Q7KVS PRELIMINARY; PRT; 8 AA.  
AC Q7KVS;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Tenascin-X (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97081760; PubMed=8923003;  
RA Speck W., Barry F., Miller W.L.;  
RT "Alternate promoters and alternate splicing of human tenascin-X, a  
RT gene with 5' and 3' ends buried in other genes."  
RL Hum. Mol. Genet. 5:1749-1758(1996).  
DR EMBL; US2700; AAC50882.1;  
FT NON TER  
SQ SEQUENCE 8 AA; 924 MW; FE2DCB46DDC76696 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 8;  
Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPREY 6  
DB 2 MPAGY 6

## RESULT 11

Q9X3K1 PRELIMINARY; PRT; 8 AA.  
AC Q9X3K1;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE Cytochrome b (Fragment).  
GN Name=petB;  
OS Prochlorococcus sp.  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=1220;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Urbach E., Chiscolm S.W.;  
RT "Genetic diversity in Prochlorococcus populations flow cytometrically  
RT sorted from the Sargasso Sea and Gulf Stream."  
RL Limnol. Oceanogr. 43:1615-1630(1998).  
DR EMBL; AF070193; AAD33233.1;  
FT NON TER  
SQ SEQUENCE 8 AA; 799 MW; 10376865B72866D3 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GPL 9  
DB 6 GPL 8

## RESULT 12

Q9TWV0 PRELIMINARY; PRT; 9 AA.  
AC Q9TWV0;

DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE Antho-RPAMIDE-NEUROPEPTIDE.

OS Anthopleura elegantissima (Sea anemone).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;  
OC Nymanthea; Actiniidae; Anthopleura.  
OX NCBI\_TaxID=6110;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93126143; PubMed=1480510;  
RA Carstensen K., Rinehart K.L., McFarlane I.D., Grimmelikhuijzen C.J.;  
RT "Isolation of Leu-Pro-Gly-Pro-Leu-Pro-Arg-Pro-NH<sub>2</sub> (Antho-RPamide),  
RT an N-terminally protected, biologically active neuropeptide from sea  
RT anemones."  
RL Peptides 13:851-857(1992).  
SQ SEQUENCE 9 AA; 943 MW; 290817637686777 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GPL 9  
DB 4 GPL 6

## RESULT 13

Q6TVP0 PRELIMINARY; PRT; 10 AA.  
AC Q6TVP0;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE PetB (Fragment).  
GN Name=petB;  
OS Wollastonia biflora (Beach sunflower) (Wedelia biflora).  
OC Chloroplast  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae;  
OX NCBI\_TaxID=101473;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dias de Moraes M., Panero J.L., Semir J.;  
RT Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY297619; ABK14772.1;  
DR GO; GO:0009507; Chloroplast; IEA.  
KM Chloroplast.  
FT NON TER  
SQ SEQUENCE 10 AA; 1068 MW; 57F74565B05866D3 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.9e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GPL 9  
DB 8 GPL 10

## RESULT 14

Q6TVP3 PRELIMINARY; PRT; 10 AA.  
AC Q6TVP3;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE PetB (Fragment).  
GN Name=petB;  
OS Otopappus epaleaceus.  
OG Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae;  
 OC Octopappus.  
 OK NCBI\_TaxID=243770;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dias de Moraes M., Panero J.L., Semir J.  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY297617; ARI14764.1; -  
 DR GO: GO:0009507; C:chloroplast; ISA.  
 KM Chloroplast.  
 FT NON TER  
 SQ SEQUENCE 10 AA; 1068 MW; 57F74565B05866D3 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+04;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 7 GPL 9  
 |||  
 Db 8 GPL 10

RESULT 15  
 O6JVP6 PRELIMINARY; PRT; 10 AA.  
 AC O6JVP6;  
 DT 05-JUL-2004 (TREMURel. 27, Created)  
 DT 05-JUL-2004 (TREMURel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMURel. 27, Last annotation update)  
 DE Pet3 (Fragment).  
 GN Name=petB;  
 OS Elaphandra paucipunctata.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae;  
 OC Elaphandra  
 OK NCBI\_TaxID=243766;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dias de Moraes M., Panero J.L., Semir J.  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY297615; ARI14764.1; -  
 DR GO: GO:0009507; C:chloroplast; IEA.  
 KW Chloroplast.  
 FT NON TER  
 SQ SEQUENCE 10 AA; 1068 MW; 57F74565B05866D3 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+04;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GPL 9  
 |||  
 Db 8 GPL 10

Search completed: November 30, 2004, 08:13:42  
 Job time : 80.3333 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:34:52 ; Search time 55.3333 Seconds  
(without alignments)  
64,831 Million cell updates/sec

Title: US-10-008-377A-6  
Perfect score: 52  
Sequence: 1 ALPIEXGPLV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 444336

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Geneseqp23sep04:\*  
2: geneseqp1980s:\*  
3: geneseqp1990s:\*  
4: geneseqp2000s:\*  
5: geneseqp2001s:\*  
6: geneseqp2002s:\*  
7: geneseqp2003as:\*  
8: geneseqp2003ds:\*  
9: geneseqp2004s:\*  
10: geneseqp2005s:\*  
11: geneseqp2006s:\*  
12: geneseqp2007s:\*  
13: geneseqp2008s:\*  
14: geneseqp2009s:\*  
15: geneseqp2010s:\*  
16: geneseqp2011s:\*  
17: geneseqp2012s:\*  
18: geneseqp2013s:\*  
19: geneseqp2014s:\*  
20: geneseqp2015s:\*  
21: geneseqp2016s:\*  
22: geneseqp2017s:\*  
23: geneseqp2018s:\*  
24: geneseqp2019s:\*  
25: geneseqp2020s:\*  
26: geneseqp2021s:\*  
27: geneseqp2022s:\*  
28: geneseqp2023s:\*  
29: geneseqp2024s:\*  
30: geneseqp2025s:\*  
31: geneseqp2026s:\*  
32: geneseqp2027s:\*  
33: geneseqp2028s:\*  
34: geneseqp2029s:\*  
35: geneseqp2030s:\*  
36: geneseqp2031s:\*  
37: geneseqp2032s:\*  
38: geneseqp2033s:\*  
39: geneseqp2034s:\*  
40: geneseqp2035s:\*  
41: geneseqp2036s:\*  
42: geneseqp2037s:\*  
43: geneseqp2038s:\*  
44: geneseqp2039s:\*  
45: geneseqp2040s:\*

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	10	AAO22988	AAO22988 Human p45
2	34	65.4	8	AAW31323	AAW31323 Dictyocaul
3	34	65.4	10	AAV04480	AAV04480 cis-epoxy
4	29	55.8	10	ADM41407	ADM41407 Human bet
5	28	53.8	10	AAO9769	AAO9769 IGFBP-2 I
6	27	51.9	8	AAW81470	AAW81470 Human cor
7	27	51.9	8	AAW81470	AAW81470 Human cor
8	26	50.0	8	AAW81470	AAW81470 Human cor
9	26	50.0	8	AAW81470	AAW81470 Human cor
10	26	50.0	8	AAW81470	AAW81470 Human cor
11	26	50.0	8	AAW81470	AAW81470 Human cor
12	26	50.0	8	AAW81470	AAW81470 Human cor
13	26	50.0	8	AAW81470	AAW81470 Human cor
14	26	50.0	8	AAW81470	AAW81470 Human cor
15	26	50.0	8	AAW81470	AAW81470 Human cor
16	26	50.0	8	AAW81470	AAW81470 Human cor
17	26	50.0	8	AAW81470	AAW81470 Human cor
18	26	50.0	8	AAW81470	AAW81470 Human cor
19	26	50.0	8	AAW81470	AAW81470 Human cor
20	26	50.0	8	AAW81470	AAW81470 Human cor
21	26	50.0	8	AAW81470	AAW81470 Human cor
22	26	50.0	8	AAW81470	AAW81470 Human cor
23	26	50.0	8	AAW81470	AAW81470 Human cor
24	26	50.0	8	AAW81470	AAW81470 Human cor
25	26	50.0	8	AAW81470	AAW81470 Human cor

26	26	50.0	9	ABJ60273	ABJ60273 184P1E2-x
27	26	50.0	9	ABJ58282	ABJ58282 184P1E2-x
28	26	50.0	9	ABJ58507	ABJ58507 184P1E2-x
29	26	50.0	9	ABJ58111	ABJ58111 184P1E2-x
30	26	50.0	10	AAW70194	AAW70194 N-termina
31	26	50.0	10	AAW47586	AAW47586 Ag85 comp
32	26	50.0	10	AAW05660	AAW05660 N-termina
33	26	50.0	10	AAW05656	AAW05656 N-termina
34	26	50.0	10	AAW05656	AAW05656 N-termina
35	26	50.0	10	AAW05656	AAW05656 N-termina
36	26	50.0	10	AAW05656	AAW05656 N-termina
37	26	50.0	10	AAW05656	AAW05656 N-termina
38	26	50.0	10	AAW05656	AAW05656 N-termina
39	26	50.0	10	AAW05656	AAW05656 N-termina
40	26	50.0	10	AAW05656	AAW05656 N-termina
41	26	50.0	10	AAW05656	AAW05656 N-termina
42	26	50.0	10	AAW05656	AAW05656 N-termina
43	26	50.0	10	AAW05656	AAW05656 N-termina
44	26	50.0	10	AAW05656	AAW05656 N-termina
45	26	50.0	10	AAW05656	AAW05656 N-termina

## ALIGNMENTS

RESULT 1  
AAO22988  
ID AAO22988 standard; peptide; 10 AA.  
XX  
AC AAO22988;  
XX  
DT 17-SEP-2003 (first entry)  
XX  
DE Human p456-65 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.  
XX  
KW HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
XX cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;  
XX oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
XX t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
XX p456-65.  
XX  
OS Homo sapiens.  
XX  
PN WO2003042243-A2.  
XX  
PD 22-MAY-2003.  
XX  
PF 14-NOV-2002; 2002WO-EP012764.  
XX  
XX 15-NOV-2001; 2001US-00008377.  
XX  
PA (MAST-) IST NAZ STUDIO E CURA DEI TUMORI.  
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX  
PI Gambacorti-Passerini C, Pasoni L;  
XX  
DR MPI; 2003-441791/41.  
XX  
PT New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
XX useful for preparing a composition for treating ALK-positive lymphoma,  
XX neuroblastoma or ALK-expressing neoplasia.  
XX  
PS Claim 1; Page 7; 33pp; English.  
XX  
CC The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma  
CC kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a  
CC cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase  
CC which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase  
CC fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell  
CC lines. More than 50% of ALCL cases possess a t(2;5) chromosomal  
CC translocation that leads to the expression of the NPM/ALK fusion protein  
CC which forms a potent oncogene when constitutively activated. Translocated  
CC ALK is a widely expressed tumour-associated antigen characteristic of ALK

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPW/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p456-65  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 XX

SO Sequence 10 AA;

Query Match 100.0%; Score 52; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0064;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10  
 |||||  
 Db 1 ALPIEYGPLV 10

RESULT 2  
 AAW31323  
 ID AAW31323 standard; peptide; 8 AA.

AC AAW31323;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 13-MAR-1998 (first entry)

DE Dictyocaulus viviparus DV18 peptide fragment 3.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity; dictyocauliasis;  
 KW cattle; immunosassay.

OS Dictyocaulus viviparus.

PN EP785253-A1.

PD 23-JUL-1997.

PF 27-DEC-1996; 96EP-00120947.

PR 19-JAN-1996; 96DE-01001754.

PA (FARH) HOECHST AG.

PI Hofmann J, Schmid K;

DR WPI; 1997-365928/34.

PT Immunogenic Dictyocaulus viviparus lungworm protein - for use in vaccines  
 PT and immunosassays.

PS Example 10; Page 6; 17pp; German.

XX This sequence represents a peptide fragment of the immunogenic protein  
 CC DV18 isolated from adult lungworms. The N-terminal of this amino acid can  
 CC be joined to the amino acids Asp, Asn, Ser, Gly or Arg. DV18 can be used  
 CC in vaccines for immunising cattle against dictyocauliasis and in an ELISA  
 CC immunosassay for determining DV18-specific antibodies in the blood of  
 CC cattle. (Updated on 25-MAR-2003 to correct PR field.)  
 XX

SO Sequence 8 AA;

Query Match 65.4%; Score 34; DB 2; Length 8;  
 Best Local Similarity 85.7%; Pred. No. 1.7e+06;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIEYGP 8  
 |||||  
 Db 2 LPIEYNP 8

RESULT 3  
 AAY04480  
 ID AAY04480 standard; peptide; 10 AA.

AC AAY04480;

DT 30-JUN-1999 (first entry)

DE cis-epoxysuccinate hydrolase internal peptide #2.

XX Rhodococcus rhodochrous; cis-epoxysuccinate hydrolase; microbial;  
 KW epoxide hydrolase; L(+); tartaric acid; food industry; additive;  
 KW soft drink; food preservative; emulsifier.

OS Rhodococcus rhodochrous.

PN EP911392-A1.

PD 28-APR-1999.

PF 24-OCT-1997; 97EP-00870168.

PR 24-OCT-1997; 97EP-00870168.

PA (PURA-) PURATOS NV.

PI Davylin T, Deslee P;

DR WPI; 1999-246412/21.

PT New epoxide hydrolase from Rhodococcus rhodochrous useful for the  
 PT production of L(+); tartaric acid.

PS Example 2; Page 7; 21pp; English.

XX The present sequence is a peptide from Rhodococcus rhodochrous IMGP-18079  
 CC cis-epoxysuccinate hydrolase (epoxide hydrolase). Epoxides are used as  
 CC chiral building blocks in the organic synthesis of fine chemicals,  
 CC especially enantiomerically pure compounds. The epoxide hydrolase can be  
 CC used to hydrolyze an epoxide, especially cis-epoxysuccinate, leading to  
 CC the production of L(+); tartaric acid. This tartaric acid is used in the  
 CC food industry as an additive in soft drinks, a food preservative and a  
 CC raw material for the synthesis of emulsifiers. L(+); tartaric acid for use  
 CC in the food industry can be produced from maleic acid but a mixture of  
 CC both the L(+) and the D(+) forms results and the D(+) form is considered  
 CC harmful to humans. L(+); tartaric acid can also be produced as a by-  
 CC product of wine fermentation but the supply is variable and dependent on  
 CC the climate from year to year. However epoxide hydrolases can be used to  
 CC produce pure L(+); tartaric acid which can therefore be used in the food  
 CC industry. Epoxide hydrolases have been identified in mammals, plants and  
 CC insects but only low levels can be obtained from these sources so are not  
 CC useful for large-scale processes. However, production of enzymes from  
 CC microorganisms culture as claimed for the new enzyme can be applied to  
 CC industrial-scale production  
 XX

SO Sequence 10 AA;

Query Match 65.4%; Score 34; DB 2; Length 10;  
 Best Local Similarity 83.3%; Pred. No. 15;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PLEYGP 8  
 |||||  
 Db 2 PLEYGP 7

RESULT 4  
 ADM41407  
 ID ADM41407 standard; peptide; 10 AA.

AC ADM41407;

XX



DT	03-JUN-2004	(first entry)
XX		
DE	Human beta3 adrenergic receptor loop 2-derived peptide R013L103.	
XX		
KM	Beta3 adrenergic receptor; G-protein coupled receptor; receptor;	
KM	angiogenesis; antihypertensive; cardiovascular; heart failure; hypertension; heart failure;	
KM	cardioprotective; vasodilator; neurotrophic; neuroprotective; nephroprotective;	
KM	cytosolic; antidiabetic; antidiabetic; immunomodulator;	
XX	signal transduction; human.	
OS	Homo sapiens.	
XX		
FH	Location/Qualifiers	
FT	Key	
FT	Modified-site	1
XX	/note= "Myristoylated"	
PN	W02004022576-A2.	
XX		
PD	18-MAR-2004.	
XX		
PF	03-SEP-2003; 2003WO-US027332.	
XX		
PR	03-SEP-2002; 2002US-0407290P.	
XX		
PA	(CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.	
PI	Ben-Sasson S, Reuveni H.	
DR	WPI; 2004-248441/23.	
PT	Identifying candidate compounds for modulating signal transduction	
PT	associated with a 7TM receptor comprises identifying a peptide region in	
PT	the 7TM receptor (unique region) and synthesizing and testing the	
PT	compound.	
XX		
PS	Claim 18; SEQ ID NO 9; 77pp; English.	
XX		
CC	The present sequence is that of R013L103, a peptide derived from the loop	
CC	2 region ADM41404 of the human beta3 adrenergic receptor, a 7-	
CC	transmembrane (7TM) receptor. The invention is based on the discovery	
CC	that short peptides from the second loop of 7TM receptors are capable of	
CC	selectively modulating signal transduction mediated by that 7TM receptor.	
CC	Methods are provided for identifying candidate compounds for the	
CC	modulation of signal transduction associated with a 7TM receptor. Such	
CC	compounds may comprise the present sequence. They can be used to treat a	
CC	disease associated with signal transduction from a 7TM receptor,	
CC	including hypertension, stroke, heart failure, neurodegenerative diseases	
CC	(including Alzheimer's disease), renal disease, psychiatric disease,	
CC	cancer, asthma, diabetes and immune disorders (all claimed). In an	
CC	example from the invention, R013L103 was shown not to increase cAMP	
CC	production from beta2-adrenergic transfected HEK293 and B16 cells	
CC	compared to the control level, but increased cAMP production 18-fold	
CC	after 120 minutes incubation with 3TFL adipocyte cells. It did not	
CC	induce melanogenesis from B16 cells.	
XX		
SQ	Sequence 10 AA:	
XX		
Query Match	55.8%; Score 28; DB 8; Length 10;	
Best Local Similarity	62.5%; Pred No. 1.2e+02;	
Matches	5; Conservative 1; Mismatches 2; Indels 0; Gaps 0	
CY	3 PIRYGPIV 10	
DB	2 PLRYGALV 9	
XX		
RESULT 5		
ID	AAB09769 standard; peptide; 10 AA.	
AC	AAB09769;	
XX		
DT	01-SEP-2000 (first entry)	

DE IGFBP-2 IGF binding domain derived amino acid sequence SEQ ID NO:138.  
XX  
XX Human; insulin like growth factor; IGF; IGFBP; binding domain;  
KM insulin like growth factor binding protein; diabetic complication;  
KM ischaemic injury; antagonist; modification; gene therapy; cyostatic;  
KM vasotropic; antidiabetic; antiparkinsonian; neuroprotective; osteopathic;  
KM antiallergic; vulnery; tranquilliser; neurologic disease; head trauma;  
KM Parkinson's disease; amyotrophic lateral sclerosis; multiple sclerosis;  
XX osteoporosis; arthritis.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200023469-A2.  
PN  
XX  
XX 27-APR-2000.  
PD  
XX  
XX 14-OCT-1999; 99NO-US023839.  
FF  
XX  
XX 16-OCT-1998; 98US-0104528P.  
PR  
XX  
XX (MUSC-) MUSC FOUND RES DEV.  
PA  
XX  
XX Rosenzweig SA, Horney MJ;  
PI  
XX  
XX WPI; 2000-339652/29.  
DR  
XX  
XX New isolated peptide having an insulin-like growth factor domain of an  
FT insulin-like growth factor binding protein, useful for treating or  
FT preventing cancer or diabetic complications, or for treating ischemic  
FT injury.  
XX  
XX  
XX Claim 4; Page 14; 106pp; English.  
PS  
XX  
XX The present invention describes an isolated peptide (A) comprising an  
XX insulin like growth factor (IGF) binding domain of an IGF-binding protein  
XX (IGFBP) or its modification. (A) binds IGF with at least the same binding  
XX affinity as the full length IGFBP. A peptide from the present invention  
XX can have cyostatic, vasotropic, antidiabetic, antiparkinsonian,  
XX neuroprotective, osteopathic, antiallergic, vulnery and tranquilliser  
XX activities. The peptide is an IGF inhibitor, IGF antagonist and can be  
XX used in gene therapy. The peptide and antagonists from the present  
XX invention are useful for the treatment or prevention of cancer or  
XX diabetic complications, and for treating ischaemic injury. Other diseases  
XX or injuries that can be treated with the fragment or antagonist include  
XX neurologic diseases and injuries, e.g. Parkinson's disease, amyotrophic  
XX lateral sclerosis, head trauma or multiple sclerosis, osteoporosis or  
XX arthritis. The biotinylated IGF is useful in therapeutic assays for IGFBP  
XX and in screening for IGFBP-mimetics (e.g. IGF antagonists). AAB09616 to  
XX AAB09773 represent amino acid sequences used in the exemplification of  
XX the present invention.  
SQ  
XX  
XX Sequence 10 AA:  
Query Match 53.8%; Score 28; DB 3; Length 10;  
Best Local Similarity 75.0%; Pred.No. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
CY 2 LPFVGGPL 9  
|||  
1 LPDERGFL 8  
DB  
RESULT 6  
AAAY8-470  
ID AA81470 standard; peptide; 10 AA.  
XX  
XX AC AA81470;  
XX  
XX 03-JUL-2000 (first entry)  
XX  
XX Human corticotatin precursor N-terminal peptide.  
DE

KW Corticosterin precursor; human; Edman degradation; N-terminal;  
 KM polypeptide sequencing apparatus; protein sequence analysis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200012994-A1.  
 XX  
 PD 09-MAR-2000.  
 XX  
 PF 26-AUG-1999; 99WO-US019460.  
 XX  
 PR 31-AUG-1998; 98US-0038401P.  
 XX  
 PA (GENTH ) GENENTECH INC.  
 XX  
 PI Henzel WJ;  
 DR WPI; 2000-237934/20.  
 XX  
 PT Apparatus for protein or polypeptide sample sequencing, uses single  
 PT reagent delivery device to supply chemicals to series of sample  
 PT cartridges.  
 XX  
 PS Example 4; Page 36; 66pp; English.  
 XX  
 CC The invention relates to a novel apparatus for polypeptide sequencing  
 CC which uses a single reagent delivery device to supply chemicals to a  
 CC series of sample cartridges. Each sample cartridge has an inlet and an  
 CC outlet. It is held horizontally in a holder at one of a series of spaced  
 CC holding stations, each including a heater. A valve selectively directs  
 CC chemicals from a single reagent delivery device to sample cartridges. The  
 CC valve also provides connections so that residues from each cartridge can  
 CC be delivered to an analyzer. The apparatus preferably uses a rapid,  
 CC chromatographic separation of amino acid derivatives to maximize sample  
 CC analysis throughput. Rapid and convenient sequencing is achieved using an  
 CC automated and highly accurate device. A single reagent delivery device  
 CC supplies chemicals to a number of sample cartridges. This reduces the  
 CC cost of manufacture and of use. Sequences AA81463-Y81475 represent the N  
 CC -terminal sequences of a variety of proteins obtained using the apparatus  
 CC of the invention  
 CC  
 SQ Sequence 10 AA;  
 Query Match 53.8%; Score 28; DB 3; Length 10;  
 Best Local Similarity 71.4%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LPIEYGP 8  
 DB 1 LPIEGGP 7  
 RESULT 7  
 AA898051  
 ID AA898051 standard; peptide; 8 AA.  
 XX  
 AC AA898051;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 01-AUG-1996 (first entry)  
 XX  
 DE Pseudosterallin G, a cyclic octapeptide.  
 XX  
 KM tyrosinase inhibitor; melanin; skin cosmetic.  
 OS Caryophyllaceae sp.  
 XX  
 XX Key Location/Qualifiers  
 FH 1  
 FT Modified-site  
 FT /note= "not an N-terminal; this amino acid condenses with  
 FT the C-terminal to form a cyclic peptide"  
 FT Modified-site 8

FT /note= "not a C-terminal; see above"  
 XX  
 PN JP07324095-A.  
 XX  
 PD 12-DEC-1995.  
 XX  
 PF 30-MAY-1994; 94JP-00116499.  
 XX  
 PR 30-MAY-1994; 94JP-00116499.  
 XX  
 PA (ITOG/) ITOGAMA H.  
 PA (HONS ) YAKULT HONSHA KK.  
 XX  
 DR WPI; 1996-065475/07.  
 XX  
 PT New cyclic peptide(s) designated pseudosterallin A-G - are tyrosinase  
 PT inhibitors and melanin formation inhibitors useful in skin cosmetics.  
 XX  
 PS Claim 7; Page 2; 13pp; Japanese.  
 XX  
 CC The cyclic peptide is one component of the peptides designated  
 CC pseudosterallins A-G which have been extracted from *Raishijin* root. These  
 CC peptides are tyrosinase inhibitors and melanin formation inhibitors which  
 CC are useful in skin cosmetics. Pseudosterallins A, B, C, D, E, F and G  
 CC demonstrated tyrosinase inhibiting IC50 values of 131, 187, 63, 100, 175,  
 CC 50 and 75 microm resp., compared with a value of 1.2 mM for arbutin.  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 CC  
 SQ Sequence 9 AA;  
 Query Match 51.9%; Score 27; DB 2; Length 8;  
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 PLEYGPL 9  
 DB 1 PPSGPL 7  
 RESULT 8  
 AA847584  
 ID AA847584 standard; peptide; 8 AA.  
 XX  
 AC AA847584;  
 XX  
 DT 13-DEC-2001 (first entry)  
 DE Ag85 complex derived peptide #15.  
 XX  
 KM Antigen 85; CD8 T-cell; vaccine; mycobacterium; detection; epitope;  
 KM lysis.  
 OS Mycobacterium tuberculosis.  
 XX  
 PN WO200170991-A1.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 20-MAR-2001; 2001WO-GB001210.  
 XX  
 PR 20-MAR-2000; 2000GB-00006693.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Klein MR;  
 DR WPI; 2001-611506/70.  
 XX  
 XX Use of a polypeptide comprising CD8 T-cell groups derived from  
 PT Mycobacterium tuberculosis antigen 85 in manufacture of a medicament for  
 PT vaccinating prophylactically or therapeutically against mycobacterial  
 PT infection.  
 XX



CC derived MHC class I (CTL) epitopes which are claimed for use in the  
 CC present invention. The invention relates to a novel expression vector  
 CC comprising a promoter operably linked to a fusion gene encoding a major  
 CC histocompatibility complex (MHC) targeting sequence, and two or more  
 CC heterologous peptide epitopes. The MHC targeting sequence may be a class  
 CC I targeting sequence, which directs an MHC class I epitope to a  
 CC cytosolic pathway or to the endoplasmic reticulum, or an MHC class II  
 CC targeting sequence, which directs extracellular antigens to enter the  
 CC endocytic pathway to be processed into antigen peptides for presentation  
 CC on MHC class II molecules. The heterologous epitopes may comprise either  
 CC helper T lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL)  
 CC epitope and a universal HTL epitope such as a pan DR epitope (PADAP). The  
 CC vectors are useful for stimulating an immune response in vivo, as well as  
 CC for use in assaying the human immunogenicity of a human T cell peptide  
 CC epitope in vivo in a non-human mammal. They provide a nucleic acid  
 CC vaccine for enhancing immunity against infectious pathogens, such as  
 CC viruses (e.g., HIV, hepatitis B (HBV) and hepatitis C (HCV)) bacteria,  
 CC protozoa (e.g., Plasmodium falciparum, the cause of malaria) and also  
 CC tumour cells and autoimmune diseases. Universal MHC class II epitopes are  
 CC advantageously combined with other MHC class I and class II epitopes to  
 CC increase the number of cells that are activated in response to a given  
 CC antigen and provide a broader population coverage of MHC-reactive alleles  
 XX

Sequence 9 AA;

Query Match 50.0%; Score 26; DB 3; Length 9;  
 Best Local Similarity 37.5%; Pred. No. 1.7e+06;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PNEYGPLV 10  
 |||  
 DB 2 PVHAGPIL 9

RESULT 11  
 AAB82787  
 ID AAB82787 standard; peptide; 9 AA.

XX AAB82787;

DT 29-OCT-2001 (first entry)

XX Epitope of mycobacterial antigen 85A.

XX Mycobacterial antigen 85A; Ag85A; infection; vaccine; diagnosis; therapy.

XX Mycobacterium tuberculosis.

XX WO200158461-A1.

XX 16-AUG-2001.

XX 12-FEB-2001; 2001WO-GB0000561.

XX 10-FEB-2000; 2000GB-00003082.

XX (GLAX) GLAXO GROUP LTD.

XX Dockrell HM, Smith SM, Brookes R;

XX WPI; 2001-536505/59.

XX Use of polypeptides comprising groups of mycobacterial antigen 85A  
 PT protein and expression vectors comprising polynucleotides encoding the  
 PT polypeptide for vaccination against Mycobacterium infection.

XX Disclosure; Page 5; 49pp; English.

XX The present sequence is that of an epitope peptide comprising amino acid  
 CC residues 48-56 of Mycobacterium tuberculosis antigen 85A (Ag85A).  
 CC Epitopes, including the present sequence, have been found in Ag85A which  
 CC cause the generation of a strong CD8 T-cell response in humans. The CD8 T  
 CC cells which recognise the epitopes are able to lyse macrophages infected

CC with live Mycobacterium tuberculosis. A polypeptide which comprises an  
 CC epitope structure of Ag85A, or an expression vector comprising a  
 CC polynucleotide encoding such a polypeptide, is used in the manufacture of  
 CC a medicament for vaccinating prophylactically or therapeutically against  
 CC infection by a mycobacterium stimulating a CD8 T-cell response. The  
 CC polypeptide or expression vector may be within an antigen-presenting  
 CC cell. A claimed vaccine composition comprises the polypeptide or  
 CC expression vector and an adjuvant or delivery system capable of  
 CC stimulating a CD8 T-cell response. Also claimed are methods of detecting  
 CC in a population of T-cells the presence or absence of CD8 T-cells that  
 CC recognise the epitope sequence, a method of diagnosing mycobacterial  
 CC infection or of testing the effectiveness of a vaccination, and a method  
 CC of treating a mycobacterium infection by administering T-cells capable of  
 CC recognising the epitope  
 XX

Sequence 9 AA;

Query Match 50.0%; Score 26; DB 4; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPVEY 6  
 |||  
 DB 2 LPVEY 6

RESULT 12  
 AAB47585  
 ID AAB47585 standard; peptide; 9 AA.

XX AAB47585;

DT 13-DEC-2001 (first entry)

XX Ag85 complex derived peptide #16.

XX Antigen 85; CD8 T-cell; vaccine; mycobacterium; detection; epitope;  
 XX lysis.

XX Mycobacterium tuberculosis.

XX WO200170991-A1.

XX 27-SEP-2001.

XX 20-MAR-2001; 2001WO-GB001210.

XX 20-MAR-2000; 2000GB-00006693.

XX (GLAX) GLAXO GROUP LTD.

XX Klein MR;

XX WPI; 2001-611506/70.

XX Use of a polypeptide comprising CD8 T-cell groups derived from  
 PT Mycobacterium tuberculosis antigen 85 in manufacture of a medicament for  
 PT vaccinating prophylactically or therapeutically against mycobacterial  
 PT infection.

XX Disclosure; Fig 1; 56pp; English.

XX The sequences given in AAB47570-86 are polypeptides derived from M.  
 CC tuberculosis antigen 85 which can be recognized by a CD8 T-cell and which  
 CC may be used in the manufacture of agent for vaccinating prophylactically  
 CC or therapeutically against infection by a mycobacterium by stimulating a  
 CC CD8 T cell response. These polypeptides are useful for vaccinating a pre-  
 CC selected host to stimulate a CD8 T cell response against a Mycobacterial  
 CC infection. They are also useful for detecting in a population of T cells  
 CC the presence or absence of CD8 T cells that recognize an epitope  
 CC sequence, by contacting the population of cells comprising CD8 T-cells  
 CC with the peptide in question and detecting whether the CD8 T-cells  
 CC recognize the peptide, by detecting the expression of a substance by the

CC T cell which indicates that the T cell have recognized the peptide or by  
 CC detecting lysis by T cells of cells that present the peptide on their  
 CC surface, the detection of lysis indicating that the T cells have  
 CC recognized the peptide

XX Sequence 9 AA;

Query Match 50.0%; Score 26; DB 4; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPEY 6  
 |||||  
 Db 4 LPVEY 8

RESULT 13  
 AAU98288  
 ID AAU98288 standard; peptide; 9 AA.

XX AAU98288;

DT 15-AUG-2002 (first entry)

DE M. tuberculosis Ag85 derived immunodominant T cell epitope #15.  
 XX Ag85; epitope; vaccine; tuberculosis; mycolyltransferase; HLA-A\*02001;  
 XX Mycobacterium tuberculosis; Mycobacterium leprae; Mycobacterium bovis;  
 XX Mycobacterium ulcerans; Mycobacterium avium; antibacterial;  
 XX MHC class I-restricted CD8<sup>+</sup> T cell; human leukocyte antigen.

XX Mycobacterium tuberculosis.

XX Key Location/Qualifiers

XX Region 2 /note="Amino acid contributing to the HLA-A\*02001  
 binding motif"

XX Region 9 /note="Amino acid contributing to the HLA-A\*02001  
 binding motif"

XX EPI211260-A1.

XX 05-JUN-2002.

XX 30-NOV-2000; 2000EP-00204268.

XX 30-NOV-2000; 2000EP-00204268.

XX (ZIEK-) ACAD ZIEKENHUIS LEIDEN.

XX Ottenhof THM, Geluk A;

XX WPI; 2002-464923/50.

XX Peptide of 8-11 amino acids derived from the Ag85 protein of  
 PT mycobacterium induce proliferation of MHC class I-restricted CD8<sup>+</sup> T cells  
 PT and are useful to vaccinate against infection by mycobacterium.

XX Claim 1; Page 7; 20pp; English.

XX The invention relates to peptide (P1) derived from an Ag85 protein of  
 CC mycobacterium (which is associated with mycolyltransferase activity and  
 CC is involved in cell wall synthesis), comprising 8-11 amino acids, and  
 CC capable of inducing proliferation of MHC class I-restricted CD8<sup>+</sup> T cells  
 CC in vitro. Also included are (1) a peptide comprising P1 flanked by amino  
 CC acids representing antigen processing sites; (2) a polypeptide comprising  
 CC at least two of the peptides in (1); (3) a nucleic acid encoding one of  
 CC the claimed peptides or polypeptides; (4) a vector comprising the above  
 CC nucleic acid; (5) a host cell comprising the above nucleic acid or vector  
 CC; and (6) detecting and/or enumerating CD8<sup>+</sup> T cells against  
 CC mycobacterium, comprising tetrameric complexes of MHC (major  
 CC histocompatibility group) class I and one of the claimed peptides or

CC polypeptides. The molecules of the invention are used to prepare a  
 CC vaccine against mycobacterium (e.g. Mycobacterium tuberculosis,  
 CC Mycobacterium leprae, Mycobacterium bovis, Mycobacterium ulcerans,  
 CC Mycobacterium avium) and thus protect against diseases such as  
 CC tuberculosis. The present sequence is an M. tuberculosis Ag85 derived  
 CC immunogenic peptide of the invention binding to HLA-A\*02001 (human  
 CC leukocyte antigen)

XX Sequence 9 AA;

Query Match 50.0%; Score 26; DB 5; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPEY 6  
 |||||  
 Db 2 LPVEY 6

RESULT 14  
 AB000162  
 ID AB000162 standard; peptide; 9 AA.

XX AB000162;

DT 29-AUG-2003 (revised)

DT 30-AUG-2002 (first entry)

DE HIV-1 GAG protein derived HLA antigen.

XX Human major histocompatibility complex; HLA; immunogen; cancer;  
 XX viral infection; vaccine; antigen; anti-HIV; virucide; cytostatic;  
 XX protozoacide.

XX Human immunodeficiency virus 1.

XX WO200220035-A1.

XX 14-MAR-2002.

XX 01-SEP-2000; 2000MO-US023913.

XX 01-SEP-2000; 2000MO-US023913.

XX (EPIIM-) EPIMUNE INC.

XX Sette A, Sidney J, Southwood S;

XX WPI; 2002-479452/51.

XX New immunogenic peptide with B7-1-like supermotif, useful in vaccines  
 PT against e.g. viral infection and cancer, induces a cytotoxic T cell  
 PT response.

XX Example 1; Page 29; 39pp; English.

XX The present invention relates to compositions containing immunogenic  
 CC peptides, which comprise a B7-like supermotif. The compositions can be  
 CC used, as vaccines, to treat or prevent viral diseases (hepatitis B or C,  
 CC Epstein-Barr, human immune deficiency, Lassa fever or cytomegalovirus),  
 CC cancers (e.g. of prostate, kidney or cervix, or lymphoma, where  
 CC associated with expression of p53, carcino-embryonal antigen or  
 CC Her2/neu), or malaria. The peptides are also useful as diagnostic agents,  
 CC e.g. to predict the outcome of a particular therapy and to identify  
 CC subjects at risk of developing a chronic infection. Nucleic acids  
 CC encoding the peptides can be used in DNA vaccines. The present sequence  
 CC is an immunogenic peptide of the invention. (Updated on 29-AUG-2003 to  
 CC standardise OS field)

XX Sequence 9 AA;

Query Match 50.0%; Score 26; DB 5; Length 9;  
 Best Local Similarity 37.5%; Pred. No. 1.7e+06;

Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 PIRXGPLV 10  
 | : | : | :  
 Db 2 PVHAGPIL 9

Db 2 PVHAGPIL 9  
 Search completed: November 30, 2004, 08:09:37  
 Job time : 57.333 secs

## RESULT 15

ABO01027  
 ID ABO01027 standard; peptide; 9 AA.  
 XX  
 AC ABO01027;  
 XX  
 DT 23-OCT-2003 (revised)  
 DT 05-AUG-2003 (first entry)  
 XX  
 DE B7-1-like supermotif-containing, immunogenic peptide 77.  
 XX  
 KW B7-1-like peptide; immunogen; HLA; human leukocyte antigen; virucide;  
 KW hepatocytic; antiinflammatory; anti-HIV; cytostatic; immunostimulant;  
 KW viral infection; cancer; hepatitis B virus infection; AIDS;  
 KW hepatitis C virus infection; cytomegalovirus infection; prostate cancer;  
 KW acquired immunodeficiency syndrome; renal carcinoma; cervical carcinoma;  
 KW lymphoma; condyloma acuminatum; chronic infection.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN US2002177694-A1.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 03-FEB-1998; 98US-00017743.  
 XX  
 PR 23-JAN-1996; 96US-00590298.  
 XX  
 PA (SETT/) SETT A.  
 PA (SIDN/) SIDNEY J.  
 PA (SOUT/) SOUTHWOOD S.  
 XX  
 PI Sette A, Sidney J, Southwood S;  
 XX  
 DR WPI: 2003-447321/42.  
 XX  
 PT Composition for treating and/or preventing viral infection and cancer  
 PT such as prostate cancer, comprises an immunogenic peptide having a B7-  
 PT like supermotif.  
 XX  
 PS Claim 1; Page 11; 14pp; English.  
 XX  
 CC The invention relates to a composition comprising an immunogenic peptide  
 CC (IP) having a B7-1-like supermotif, where IP is selected from one of the  
 CC 127 peptide sequences (appearing as ABO00951-ABO01082). The composition  
 CC is useful for treating and/or preventing viral infection and cancer.  
 CC including infection by hepatitis B, hepatitis C or cytomegalovirus (CMV),  
 CC prostate cancer, acquired immunodeficiency syndrome (AIDS), renal  
 CC carcinoma, cervical carcinoma, lymphoma and condyloma acuminatum. The  
 CC composition is also useful for eliciting an immune response against a  
 CC desired antigen, in therapeutic or diagnostic applications, for  
 CC preventing the evolution from acute to chronic infection, for the  
 CC treatment of chronic infection, and to stimulate the immune system to  
 CC eliminate virus-infected cells in carriers. The peptides are HLA (human  
 CC leukocyte antigen) binding peptides. The present sequence is an  
 CC immunogenic peptide of the invention, having a B7-1-like supermotif.  
 CC (Updated on 23-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 9 AA;

Query Match 50.0%; Score 26; DB 6; Length 9;  
 Best Local Similarity 37.5%; Pred. No. 1.7e+06;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIRXGPLV 10  
 | : | : | :  
 Db 2 PVHAGPIL 9

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 08:13:53 ; Search time 44 Seconds

(without alignments)  
80.612 Million cell updates/sec

Title: US-10-008-377A-6  
Perfect score: 52  
Sequence: 1 ALPIRYGPLY 10

Scoring table: BIOSIM62  
Gapop 10.0, Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 177356

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum March 0%  
Maximum March 100%

Listing first 45 summaries

Database :	Published Applications AA:
1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/US07_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
19:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
20:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	52	100.0	10	US-10-008-377-6	Sequence 6, Appl1
2	27	51.9	7	US-10-041-030-14	Sequence 14, Appl1
3	26	50.0	9	US-09-017-743C-77	Sequence 77, Appl1
4	26	50.0	9	US-10-371-525-404	Sequence 404, Appl1
5	26	50.0	9	US-10-371-525-404	Sequence 404, Appl1
6	26	50.0	9	US-10-371-525-404	Sequence 404, Appl1
7	26	50.0	9	US-10-371-525-404	Sequence 404, Appl1
8	26	50.0	9	US-10-371-525-404	Sequence 404, Appl1
9	26	50.0	9	US-10-371-525-404	Sequence 404, Appl1
10	24	46.2	10	US-09-572-404B-461	Sequence 461, Appl1
11	24	46.2	10	US-09-932-165-801	Sequence 165, Appl1
12	24	46.2	10	US-08-944-824-11	Sequence 11, Appl1
13	24	46.2	10	US-09-898-860-126	Sequence 126, Appl1
14	24	46.2	10	US-09-935-476-4	Sequence 4, Appl1

14	24	46.2	10	US-10-094-699-108	Sequence 108, Appl1
15	24	46.2	10	US-10-359-431-25	Sequence 25, Appl1
16	24	46.2	10	US-10-371-525-404	Sequence 226, Appl1
17	24	46.2	10	US-10-025-066-86	Sequence 86, Appl1
18	24	46.2	10	US-10-371-069-326	Sequence 226, Appl1
19	24	46.2	10	US-10-371-645-326	Sequence 226, Appl1
20	24	46.2	10	US-10-371-937-34	Sequence 24, Appl1
21	24	46.2	10	US-10-371-280-226	Sequence 24, Appl1
22	24	46.2	10	US-10-448-581-45	Sequence 45, Appl1
23	24	46.2	10	US-10-149-138-1189	Sequence 4190, Appl1
24	24	46.2	10	US-10-149-138-1190	Sequence 4190, Appl1
25	24	46.2	10	US-10-149-138-1191	Sequence 4192, Appl1
26	24	46.2	10	US-10-149-138-1192	Sequence 4192, Appl1
27	24	46.2	10	US-10-149-138-1193	Sequence 4193, Appl1
28	24	46.2	10	US-10-149-138-1194	Sequence 4194, Appl1
29	24	46.2	10	US-10-149-138-1195	Sequence 4195, Appl1
30	24	46.2	10	US-10-149-138-1196	Sequence 4196, Appl1
31	24	46.2	10	US-10-149-138-1197	Sequence 4197, Appl1
32	24	46.2	10	US-10-149-138-1198	Sequence 4198, Appl1
33	24	46.2	10	US-10-149-138-1199	Sequence 4199, Appl1
34	24	46.2	10	US-10-149-138-1200	Sequence 4200, Appl1
35	24	46.2	10	US-10-149-135-2176	Sequence 2176, Appl1
36	24	46.2	10	US-10-149-135-2177	Sequence 2177, Appl1
37	24	46.2	10	US-10-149-135-2178	Sequence 2178, Appl1
38	24	46.2	10	US-10-149-135-2179	Sequence 2179, Appl1
39	24	46.2	10	US-10-149-135-2180	Sequence 2180, Appl1
40	24	46.2	10	US-10-149-135-2181	Sequence 2181, Appl1
41	24	46.2	10	US-10-149-135-2182	Sequence 2182, Appl1
42	24	46.2	10	US-10-149-135-2183	Sequence 2183, Appl1
43	24	46.2	10	US-10-149-135-2184	Sequence 2184, Appl1
44	24	46.2	10	US-10-149-135-2185	Sequence 2185, Appl1
45	24	46.2	10	US-10-619-454-1	Sequence 1, Appl1

## ALIGNMENTS

RESULT 1  
US-10-008-377-6  
Sequence 6, Application US/10008377  
Publication No. US20030157101A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
TITLE OF INVENTION: Immunochemical ALK Peptides  
FILE REFERENCE: 045922/241203  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 6  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-6

Query Match  
Best Local Similarity 100.0%; Pred. No. 0.0098;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIRYGPLY 10  
DB 1 ALPIRYGPLY 10

US-10-041-030-14  
Sequence 14, Application US/10041030  
Publication No. US20020150934A1  
GENERAL INFORMATION:  
APPLICANT: Powers, Scott  
APPLICANT: Mu, David  
APPLICANT: Xiang, Phil

APPLICANT: Peng, Yue  
APPLICANT: Tularex Inc.  
TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian  
TITLE OF INVENTION: pellino polypeptides and polynucleotides  
FILE REFERENCE: 018781-006810US  
CURRENT APPLICATION NUMBER: US/10/041,030  
CURRENT FILING DATE: 2001-12-28  
PRIOR APPLICATION NUMBER: US 60/259,502  
PRIOR FILING DATE: 2001-01-02  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO: 14  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: amino acid  
OTHER INFORMATION: sequence identity in comparison of pellino 1 and  
US-10-041-030-14

Query Match  
Best Local Similarity 51.9%; Score 27; DB 13; Length 7;  
Pred. No. 1,4e+06;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 PIRGAPL 9  
DB 1 PIRGAPL 7

RESULT 3  
US-09-017-743C-77  
Sequence 77, Application US/09017743C  
Patent No. US2002017694A1  
GENERAL INFORMATION:  
APPLICANT: Sette, Alessandro  
Sidney, John  
Southwood, Scott  
TITLE OF INVENTION: HLA Binding Peptides and Their  
Uses  
NUMBER OF SEQUENCES: 146  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/017,743C  
FILING DATE: 03-Feb-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/590,298  
FILING DATE: 23-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Parent, Annette S.  
REGISTRATION NUMBER: 42,058  
REFERENCE/DOCKET NUMBER: 018623-008050US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 77:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 77:  
US-09-017-743C-77

Query Match  
Best Local Similarity 50.0%; Score 26; DB 9; Length 9;  
Pred. No. 1,4e+06;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIRGAPL 10  
DB 2 PVHAGPIL 9

RESULT 4  
US-10-371-525-404  
Sequence 404, Application US/10371525  
Publication No. US20030203869A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John D.  
APPLICANT: Hermanson, Gary G.  
APPLICANT: Sette, Alessandro  
APPLICANT: Ishioka, Glenn Y.  
APPLICANT: Livingston, Brian  
APPLICANT: Chesnut, Robert W.  
APPLICANT: EpiImmune Inc.  
TITLE OF INVENTION: Expression Vectors for Stimulating an  
Immune Response and Methods of Using the Same  
FILE REFERENCE: 39963-20022.01  
CURRENT APPLICATION NUMBER: US/10/371,525  
PRIOR FILING DATE: 2003-02-21  
PRIOR APPLICATION NUMBER: US 09/311,784  
PRIOR FILING DATE: 1999-05-13  
PRIOR APPLICATION NUMBER: US 60/085,751  
PRIOR FILING DATE: 1998-05-15  
NUMBER OF SEQ ID NOS: 463  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 404  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HIV GAG 248 (peptide 1292.14)  
US-10-371-525-404

Query Match  
Best Local Similarity 50.0%; Score 26; DB 14; Length 9;  
Pred. No. 1,4e+06;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIRGAPL 10  
DB 2 PVHAGPIL 9

RESULT 5  
US-10-371-069-404  
Sequence 404, Application US/10371069  
Publication No. US20030216342A1  
GENERAL INFORMATION:  
APPLICANT: EpiImmune Inc.  
APPLICANT: Fikes, John D.  
APPLICANT: Hermanson, Gary G.  
APPLICANT: Sette, Alessandro  
APPLICANT: Ishioka, Glenn Y.  
APPLICANT: Livingston, Brian  
APPLICANT: Chesnut, Robert W.  
APPLICANT: EpiImmune Inc.  
TITLE OF INVENTION: Expression Vectors for Stimulating an  
Immune Response and Methods of Using the Same  
FILE REFERENCE: 39963-20022.10  
CURRENT APPLICATION NUMBER: US/10/371,069  
PRIOR FILING DATE: 2003-02-21  
PRIOR APPLICATION NUMBER: US 09/078,904



```
/ PRIOR FILING DATE: 1998-05-13
/ PRIOR APPLICATION NUMBER: US 60/085,751
/ PRIOR FILING DATE: 1998-05-15
/ NUMBER OF SEQ ID NOS: 463
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 404
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: HIV GAG 248 (peptide 1292.14)
US-10-371-069-404

Query Match
Best Local Similarity 50.0%; Score 26; DB 14; Length 9;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIEVGLV 10
DB 2 PVHAGPII 9

RESULT 6
US-10-371-645-404
/ Sequence 404, Application US/103771645
/ Publication No. US20030216343A1
/ GENERAL INFORMATION:
/ APPLICANT: EPIMUNE INC.
/ APPLICANT: Fikes, John D.
/ APPLICANT: Hermanson, Gary G.
/ APPLICANT: Settle, Alessandro
/ APPLICANT: Ishioaka, Glenn Y.
/ APPLICANT: Livingston, Brian
/ APPLICANT: Chesnut, Robert W.
/ APPLICANT: Epimune Inc.
/ TITLE OF INVENTION: Expression Vectors for Stimulating an
/ TITLE OF INVENTION: Immune Response and Methods of Using the Same
/ FILE REFERENCE: 39963-20022.11
/ CURRENT FILING DATE: 2003-06-20
/ PRIOR APPLICATION NUMBER: US 09/078,904
/ PRIOR FILING DATE: 1998-05-13
/ PRIOR APPLICATION NUMBER: US 60/085,751
/ PRIOR FILING DATE: 1998-05-15
/ NUMBER OF SEQ ID NOS: 463
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 404
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: HIV GAG 248 (peptide 1292.14)
US-10-371-645-404

Query Match
Best Local Similarity 37.5%; Score 26; DB 14; Length 9;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIEVGLV 10
DB 2 PVHAGPII 9

RESULT 7
US-10-371-260-404
/ Sequence 404, Application US/103771260
/ Publication No. US20030220285A1
/ GENERAL INFORMATION:
/ APPLICANT: EPIMUNE INC.
/ APPLICANT: Fikes, John D.
/ APPLICANT: Hermanson, Gary G.
/ APPLICANT: Settle, Alessandro
/ APPLICANT: Ishioaka, Glenn Y.
```

```
/ APPLICANT: Livingston, Brian
/ APPLICANT: Chesnut, Robert W.
/ APPLICANT: Epimune Inc.
/ TITLE OF INVENTION: Expression Vectors for Stimulating an
/ TITLE OF INVENTION: Immune Response and Methods of Using the Same
/ FILE REFERENCE: 39963-20022.13
/ CURRENT FILING DATE: 2003-02-21
/ PRIOR APPLICATION NUMBER: US 09/078,904
/ PRIOR FILING DATE: 1998-05-13
/ PRIOR APPLICATION NUMBER: US 60/085,751
/ PRIOR FILING DATE: 1998-05-15
/ NUMBER OF SEQ ID NOS: 463
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 404
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: HIV GAG 248 (peptide 1292.14)
US-10-371-260-404

Query Match
Best Local Similarity 50.0%; Score 26; DB 14; Length 9;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIEVGLV 10
DB 2 PVHAGPII 9

RESULT 8
US-10-203-562-1
/ Sequence 1, Application US/10203562
/ Publication No. US20040037837A1
/ GENERAL INFORMATION:
/ APPLICANT: HAZEL M. DOCKRELL
/ APPLICANT: STEVEN M. SMITH
/ APPLICANT: ROGER BROOKES
/ TITLE OF INVENTION: MYCOBACTERIUM AG85 COMPLEX-SPECIFIC T CELL PEPTIDES AND USE IN
/ TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC APPLICATIONS THEREOF
/ FILE REFERENCE: 117-409 / N.784688 GCM
/ CURRENT FILING DATE: 2002-08-12
/ PRIOR APPLICATION NUMBER: US/10/203,562
/ PRIOR FILING DATE: 2001-02-12
/ PRIOR APPLICATION NUMBER: GB 0003082.5
/ PRIOR FILING DATE: 2000-02-10
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO: 1
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Mycobacterium tuberculosis
US-10-203-562-1

Query Match
Best Local Similarity 80.0%; Score 26; DB 15; Length 9;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEV 6
DB 2 LPVEY 6

RESULT 9
US-09-572-404B-461
/ Sequence 461, Application US/09572404B
/ Publication No. US20030078374A1
/ GENERAL INFORMATION:
/ APPLICANT: Proteom Ltd
/ TITLE OF INVENTION: Complementary peptide ligands from the human genome
/ FILE REFERENCE: Human patent
```

CURRENT APPLICATION NUMBER: US/09/572,404B  
CURRENT FILING DATE: 2000-05-17  
NUMBER OF SEQ ID NOS: 4203  
SOFTWARE: ProPatent version 1.0  
SEQ ID NO 461  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo Sapiens  
FEATURE:  
OTHER INFORMATION: Sequence located in CLAP2 OR AP17 at 16-25 and may interact with  
US-09-572-404B-461

Query Match 48.1%; Score 25; DB 10; Length 10;  
Best Local Similarity 71.4%; Pred. No. 7.5e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPIEYGP 8  
DB 4 LPSESGP 10

RESULT 10  
US-09-932-165-801  
Sequence 801, Application US/09932165  
Publication No. US20030134784A1  
GENERAL INFORMATION:  
APPLICANT: RAITANO, ARTHUR  
APPLICANT: CHALLITH-ED, PTA M.  
APPLICANT: FARIS, MARI  
APPLICANT: SAFRAN, DOUGLAS  
APPLICANT: AFAR, DANIEL  
APPLICANT: LEVIN, ELANA  
APPLICANT: HUBERT, RENE  
APPLICANT: GE, WANGMAO  
APPLICANT: JAKBOVITS, AVA  
TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
TITLE OF INVENTION: 83P23 AND CAPTEP211 USED IN TREATMENT AND  
TITLE OF INVENTION: DETECTION OF CANCER  
FILE REFERENCE: 5138-20014.00  
CURRENT APPLICATION NUMBER: US/09/932,165  
CURRENT FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: 60/226,329  
PRIOR FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 1508  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 801  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-801

Query Match 46.2%; Score 24; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YGPL 9  
DB 2 YGPL 5

RESULT 11  
US-08-344-824-11  
Sequence 11, Application US/08344824  
Publication No. US20030152560A1  
GENERAL INFORMATION:  
APPLICANT: SETTE, Alessandro  
APPLICANT: SIDNEY, John  
TITLE OF INVENTION: H2A BINDING PEPTIDES AND THEIR USES  
NUMBER OF SEQUENCES: 399  
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Townsend and Townsend Hourie and Crew  
STREET: One Market Plaza, Stewart Street Tower, 20th  
STREET: Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/344,824  
FILING DATE: 23-NOV-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/278,634  
FILING DATE: 21-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 14137-80-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-344-824-11

Query Match 46.2%; Score 24; DB 8; Length 10;  
Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPIEYGP 10  
DB 2 LPSPYPSV 10

RESULT 12  
US-09-898-860-126  
Sequence 126, Application US/09898860  
Publication No. US20030144482A1  
GENERAL INFORMATION:  
APPLICANT: KAKAKI, YUTAKA; ROSENBERG, STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND THEIR USE IN DIAGNOSTIC AND THERAPEUTIC METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/898,860  
FILING DATE: 03-JUL-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/267,439  
FILING DATE: <Unknown>

APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPEI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 759-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 126:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 126:  
US-09-898-860-126

Query Match 46.2%; Score 24; DB 10; Length 10;  
Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPIEYGPLV 10  
DB 2 LPSDYFSPV 10

RESULT 13  
US-09-935-476-4  
Sequence 4, Application US/09935476  
Publication No. US20040096445A1  
GENERAL INFORMATION:  
APPLICANT: Epimmune, Inc.  
APPLICANT: Sidney, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Grey, Howard  
APPLICANT: Southwood, Scott  
TITLE OF INVENTION: SUBUNIT VACCINES WITH A2 SUPERMOTIFS  
FILE REFERENCE: 39963-20029.20  
CURRENT APPLICATION NUMBER: US/09/935,476  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: US 09/346,105  
PRIOR FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: US 60/264,969  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Analog of the HBV core 18027 epitope  
US-09-935-476-4  
Query Match 46.2%; Score 24; DB 11; Length 10;  
Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPIEYGPLV 10  
DB 2 LPSDYFSPV 10

RESULT 14  
US-10-094-699-108  
Sequence 108, Application US/10094699  
Publication No. US20030046714A1  
GENERAL INFORMATION:

APPLICANT: SIMARD, John, J.L.  
APPLICANT: DIAMOND, David, C.  
TITLE OF INVENTION: ANTI-NEOVASCULATURE PREPARATIONS FOR  
TITLE OF INVENTION: CANCER  
FILE REFERENCE: CTIIMM.015A  
CURRENT APPLICATION NUMBER: US/10/094,699  
CURRENT FILING DATE: 2002-06-21  
PRIOR APPLICATION NUMBER: US 60/274,063  
PRIOR FILING DATE: 2001-03-07  
NUMBER OF SEQ ID NOS: 109  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 108  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-094-699-108  
Query Match 46.2%; Score 24; DB 14; Length 10;  
Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPIEYGPLV 10  
DB 2 LPSDYFSPV 10

RESULT 15  
US-10-359-431-25  
Sequence 25, Application US/10359431  
Publication No. US20030171538A1  
GENERAL INFORMATION:  
APPLICANT: Chiesari, Francis V.  
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T  
Lymphocyte Responses to Hepatitis B Virus  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/359,431  
FILING DATE: 05-Feb-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,540  
FILING DATE: 26-AUG-1991  
APPLICATION NUMBER: US 07/935,898  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: US 08/100,870  
FILING DATE: 02-AUG-1993  
APPLICATION NUMBER: WO PCT/US94/08685  
FILING DATE: 01-AUG-1994  
APPLICATION NUMBER: US 08/591,502  
FILING DATE: 20-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Lockyer, Jean M.  
REGISTRATION NUMBER: 44,879  
REFERENCE/DOCKET NUMBER: 014740-000231US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-10-359-431-25

Query Match 46.2%; Score 24; DB 14; Length 10;  
Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LPIEXGPLV 10  
||:|  
Db 2 LPSDYFPEV 10

Search completed: November 30, 2004, 08:31:01  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:41:57 ; Search time 27 Seconds

(without alignments)  
24,562 Million cell updates/sec

Title: US-10-008-377a-6

Perfect score: 52

Sequence: 1 ADP1EXGPV 10

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 110780

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: /cgn2\_6/prodata/1/aa/5A.COMB.pep.\*  
2: /cgn2\_6/prodata/1/aa/5B.COMB.pep.\*  
3: /cgn2\_6/prodata/1/aa/6A.COMB.pep.\*  
4: /cgn2\_6/prodata/1/aa/6B.COMB.pep.\*  
5: /cgn2\_6/prodata/1/aa/6C.COMB.pep.\*  
6: /cgn2\_6/prodata/1/aa/6D.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	65.4	10	3	US-09-530-058-3
2	27	51.9	7	3	US-08-786-458-3
3	26	50.0	9	4	US-09-311-784A-404
4	26	50.0	10	2	US-08-733-505A-32
5	26	50.0	10	2	US-08-706-741B-67
6	26	50.0	10	2	US-08-924-655A-67
7	26	50.0	10	3	US-09-001-984C-92
8	26	50.0	10	3	US-09-001-984C-97
9	26	50.0	10	3	US-09-001-984C-101
10	26	50.0	10	4	US-09-396-347F-92
11	26	50.0	10	4	US-09-396-347F-97
12	26	50.0	10	4	US-09-396-347F-101
13	25	48.1	7	3	US-08-604-965E-10
14	25	48.1	7	3	US-08-604-965E-15
15	25	48.1	7	6	5258287-43
16	25	48.1	8	6	5258287-44
17	25	48.1	10	4	US-09-239-043D-2314
18	24	46.2	9	1	US-08-276-452A-58
19	24	46.2	9	1	US-08-798-744-58
20	24	46.2	10	1	US-08-195-075-12
21	24	46.2	10	1	US-08-363-691-9
22	24	46.2	10	2	US-08-417-174-126
23	24	46.2	10	3	US-08-155-359A-647
24	24	46.2	10	3	US-09-267-439-126
25	24	46.2	10	3	US-08-992-877-1
26	24	46.2	10	4	US-09-311-784A-226
27	24	46.2	10	4	US-09-073-138-126

28	24	46.2	10	4	US-09-497-232-4	Sequence 4, Appl
29	24	46.2	10	4	US-09-543-608A-43	Sequence 43, Appl
30	24	46.2	10	4	US-08-591-502B-25	Sequence 25, Appl
31	24	46.2	10	4	US-08-234-784B-97	Sequence 97, Appl
32	24	46.2	10	4	US-09-239-043D-2254	Sequence 2254, Ap
33	24	46.2	10	4	US-09-239-043D-2307	Sequence 2307, Ap
34	24	46.2	10	4	US-09-239-043D-2313	Sequence 2313, Ap
35	24	46.2	10	4	US-09-239-043D-2317	Sequence 2317, Ap
36	24	46.2	10	5	PCT-US95-01672-12	Sequence 12, Appl
37	23	44.2	8	3	US-08-444-818-417	Sequence 417, Appl
38	23	44.2	9	1	US-08-615-181-18	Sequence 18, Appl
39	23	44.2	9	4	US-09-311-784A-370	Sequence 370, Appl
40	23	44.2	9	4	US-09-820-174-33	Sequence 33, Appl
41	23	44.2	9	4	US-09-820-195A-33	Sequence 33, Appl
42	23	44.2	10	3	US-09-561-366B-33	Sequence 33, Appl
43	23	44.2	10	4	US-10-114-176-33	Sequence 33, Appl
44	23	44.2	10	4	US-09-239-043D-2288	Sequence 2288, Ap
45	23	44.2	10	4	US-09-239-043D-2308	Sequence 2308, Ap

#### ALIGNMENTS

RESULT 1  
US-09-530-058-3  
Sequence 3, Application US/09530058  
Patent No. 6379938  
GENERAL INFORMATION:

APPLICANT: <Unknown>

TITLE OF INVENTION: Epoxide hydrolase

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Knobe, Martens, Olson & Bear

STREET: 501 West Broadway  
CITY: San Diego

STATE: California  
COUNTRY: USA

ZIP: 92101-3505

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95

SOFTWARE: Word  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/530,058  
FILING DATE: 03-Jul-2000

ATTORNEY/AGENT INFORMATION:  
NAME: Harte, Daniel

REGISTRATION NUMBER: 40,637  
REFERENCE/DOCKET NUMBER: VANM150.001APC

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids

TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULAR TYPE: peptide

HYPOTHEICAL: NO  
ANTI-SENSE: NO

FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-530-058-3

Query Match 65.4%, Score 34, DB 3, Length 10,  
Best Local Similarity 83.3%, Pred. No. 5.2,  
Matches 5, Conservative 1, Mismatches 0, Gaps 0,  
Indels 0,  
CY 3 PILEUP 8  
DB 2 PILEUP 7

RESULT 2  
US-08-786-455B-3  
Sequence 3, Application US/08786455B  
Patent No. 6193971  
GENERAL INFORMATION:  
APPLICANT: HOFMANN, Joachim  
APPLICANT: SCHMID, Karlheinz  
TITLE OF INVENTION: DICTYOCALUS VIVIPARUS ANTIGEN FOR  
AND FOR VACCINATION  
TITLE OF INVENTION: DIAGNOSING LUNGWORM INFESTATION  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/786,455B  
FILING DATE: 21-JAN-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 01 754.8  
FILING DATE: 19-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: GRANADOS, Patricia D.  
REGISTRATION NUMBER: 33,683  
REFERENCE/DOCKET NUMBER: 18748/327  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-786-455B-3  
Query Match  
Best Local Similarity 51.9%; Score 27; DB 3; Length 7;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 PIRY 6  
DB 1 PIRY 5  
RESULT 3  
US-09-311-784A-404  
Sequence 404, Application US/09311784A  
Patent No. 6534482  
GENERAL INFORMATION:  
APPLICANT: Fikes, John D.  
APPLICANT: Hermanson, Gary G.  
APPLICANT: Sette, Alessandro  
APPLICANT: Ishioka, Glenn Y.  
APPLICANT: Livingston, Brian  
APPLICANT: Chesnut, Robert W.  
APPLICANT: Epiimmune, Inc.  
TITLE OF INVENTION: Expression Vectors for Stimulating an  
Immune Response and Methods of Using the Same  
FILE REFERENCE: 39963-20022.01  
CURRENT APPLICATION NUMBER: US/09/311,784A  
CURRENT FILING DATE: 1999-05-13

PRIOR APPLICATION NUMBER: US 60/085,751  
PRIOR FILING DATE: 1998-05-15  
NUMBER OF SEQ ID NOS: 463  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 404  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HIV GAG 248 (peptide 1292.14)  
US-09-311-784A-404  
Query Match  
Best Local Similarity 50.0%; Score 26; DB 4; Length 9;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 3 PIRYGLV 10  
DB 2 PIRYGLV 9  
RESULT 4  
US-08-733-505A-32  
Sequence 32, Application US/08733505A  
Patent No. 5856445  
GENERAL INFORMATION:  
APPLICANT: KORSMEYER, STANLEY J.  
TITLE OF INVENTION: SERINE SUBSTITUTED MUTANTS OF  
BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAERKAMP, L.C.  
STREET: 7733 FORSYTH BLVD., SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: USA  
ZIP: 63105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/733,505A  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 965458  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-733-505A-32  
Query Match  
Best Local Similarity 50.0%; Score 26; DB 2; Length 10;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 3 PIRYGLV 10  
DB 2 PIRYGLV 9  
RESULT 5  
US-08-706-741B-67

```

; Sequence 67, Application US/08706741B
; Patent No. 5955593
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, STANLEY J.
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: USA
; ZIP: 63146
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,741B
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 965017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-706-741B-67

Query Match
Best Local Similarity 50.0%; Score 26; DB 2; Length 10;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PLEYGPLV 10
DB 2 FMSYGRLL 9

RESULT 6
US-08-924-695A-67
; Sequence 67, Application US/08924695A
; Patent No. 5985583
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, STANLEY J.
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/924,695A
; FILING DATE: 09-SEP-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.

```

```

; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 971798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-924-695A-67

Query Match
Best Local Similarity 50.0%; Score 26; DB 2; Length 10;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PLEYGPLV 10
DB 2 FMSYGRLL 9

RESULT 7
US-09-001-984C-92
; Sequence 92, Application US/09001984C
; Patent No. 6245331
; GENERAL INFORMATION:
; APPLICANT: Laal, Susan
; APPLICANT: Zolla-Pazner, Susan
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE
; FILE REFERENCE: NYU-011
; CURRENT APPLICATION NUMBER: US/09/001,984C
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: 60/034,003
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 92
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis strain H37Rv
; US-09-001-984C-92

Query Match
Best Local Similarity 80.0%; Score 26; DB 3; Length 10;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPYEV 6
DB 6 LPYEV 10

RESULT 8
US-09-001-984C-97
; Sequence 97, Application US/09001984C
; Patent No. 6245331
; GENERAL INFORMATION:
; APPLICANT: Laal, Susan
; APPLICANT: Zolla-Pazner, Susan
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE
; FILE REFERENCE: NYU-011
; CURRENT APPLICATION NUMBER: US/09/001,984C
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: 60/034,003
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 97
; LENGTH: 10
; TYPE: PRT

```

ORGANISM: Mycobacterium tuberculosis strain H37RV  
US-09-001-984C-97

Query Match 50.0%; Score 26; DB 3; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1,4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRY 6  
||:|  
DB 6 LPEY 10

RESULT 9  
US-09-001-984C-101

; Sequence 101, Application US/09001984C  
; Patent No. 6245331  
; GENERAL INFORMATION:  
; APPLICANT: Laal, Suman  
; APPLICANT: Zolla-Pazner, Susan  
; APPLICANT: Belisle, John T  
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
; FILE REFERENCE: NYU-011  
; CURRENT APPLICATION NUMBER: US/09/001,984C  
; CURRENT FILING DATE: 1997-12-31  
; PRIOR APPLICATION NUMBER: 60/034,003  
; PRIOR FILING DATE: 1996-12-31  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 101  
; TYPE: PRT  
; LENGTH: 10  
; ORGANISM: Mycobacterium tuberculosis strain H37RV  
US-09-001-984C-101

Query Match 50.0%; Score 26; DB 3; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1,4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRY 6  
||:|  
DB 6 LPEY 10

RESULT 10  
US-09-396-347F-92  
; Sequence 92, Application US/09396347F  
; Patent No. 6506384  
; GENERAL INFORMATION:  
; APPLICANT: Laal, Suman  
; APPLICANT: Zolla-Pazner, Susan  
; APPLICANT: Belisle, John T  
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
; FILE REFERENCE: 32004-169276  
; CURRENT APPLICATION NUMBER: US/09/396,347F  
; CURRENT FILING DATE: 1999-09-14  
; PRIOR APPLICATION NUMBER: 09/001,984  
; PRIOR FILING DATE: 1997-12-31  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 92  
; TYPE: PRT  
; LENGTH: 10  
; ORGANISM: Mycobacterium tuberculosis strain H37RV  
US-09-396-347F-92

Query Match 50.0%; Score 26; DB 4; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1,4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRY 6  
||:|  
DB 6 LPEY 10

RESULT 11  
US-09-396-347F-97

; Sequence 97, Application US/09396347F  
; Patent No. 6506384  
; GENERAL INFORMATION:  
; APPLICANT: Laal, Suman  
; APPLICANT: Zolla-Pazner, Susan  
; APPLICANT: Belisle, John T  
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
; FILE REFERENCE: 32004-169276  
; CURRENT APPLICATION NUMBER: US/09/396,347F  
; CURRENT FILING DATE: 1999-09-14  
; PRIOR APPLICATION NUMBER: 09/001,984  
; PRIOR FILING DATE: 1997-12-31  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 97  
; TYPE: PRT  
; LENGTH: 10  
; ORGANISM: Mycobacterium tuberculosis strain H37RV  
US-09-396-347F-97

Query Match 50.0%; Score 26; DB 4; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1,4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRY 6  
||:|  
DB 6 LPEY 10

RESULT 12  
US-09-396-347F-101

; Sequence 101, Application US/09396347F  
; Patent No. 6506384  
; GENERAL INFORMATION:  
; APPLICANT: Laal, Suman  
; APPLICANT: Zolla-Pazner, Susan  
; APPLICANT: Belisle, John T  
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
; FILE REFERENCE: 32004-169276  
; CURRENT APPLICATION NUMBER: US/09/396,347F  
; CURRENT FILING DATE: 1999-09-14  
; PRIOR APPLICATION NUMBER: 09/001,984  
; PRIOR FILING DATE: 1997-12-31  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 101  
; TYPE: PRT  
; LENGTH: 10  
; ORGANISM: Mycobacterium tuberculosis strain H37RV  
US-09-396-347F-101

Query Match 50.0%; Score 26; DB 4; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1,4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRY 6  
||:|  
DB 6 LPEY 10

RESULT 13  
US-08-604-965E-10

; Sequence 10, Application US/08604965E  
; Patent No. 6046033  
; GENERAL INFORMATION:  
; APPLICANT: Goto, Masaki  
; APPLICANT: Tsuda, Sioke  
; APPLICANT: Yano, Kazuki  
; APPLICANT: Kobayashi, Fumie  
; APPLICANT: Yamaguchi, Kyoji



APPLICANT: Washida, Naohiro  
APPLICANT: Satake, Toshiko  
APPLICANT: Morinaga, Tomonori  
APPLICANT: Ueda, Masatsugu  
APPLICANT: Higashio, Kanji  
TITLE OF INVENTION: Basic Osteoblast Growth Factor II (bOGF)  
TITLE OF INVENTION: II  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burgess, Ryan and Wayne  
STREET: 370 Lexington Avenue, Suite 2105  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10017  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5 1/4 inch diskette  
COMPUTER: PC/S LIMITED SYSTEM 200  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/604,965E  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/01270  
FILING DATE: June 27, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Wayne, Milton J.  
REGISTRATION NUMBER: 17,906  
REFERENCE/DOCKET NUMBER: U-WP-5212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-683-8150  
TELEFAX: 212-532-4285  
TELEX: 423794  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-604-965E-10

Query Match 48.1%; Score 25; DB 3; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EYGP 8  
|||  
Db 3 EYGP 6

RESULT 14  
US-08-604-965E-15  
Sequence 15, Application US/08604965E  
Patent No. 6046033  
GENERAL INFORMATION:  
APPLICANT: Goto, Masaaki  
APPLICANT: Tsuda, Eisuke  
APPLICANT: Yano, Kazuki  
APPLICANT: Kobayashi, Fumie  
APPLICANT: Yamaguchi, Kyoji  
APPLICANT: Washida, Naohiro  
APPLICANT: Satake, Toshiko  
APPLICANT: Morinaga, Tomonori  
APPLICANT: Ueda, Masatsugu  
APPLICANT: Higashio, Kanji  
TITLE OF INVENTION: Basic Osteoblast Growth Factor II (bOGF)  
TITLE OF INVENTION: II  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burgess, Ryan and Wayne

STREET: 370 Lexington Avenue, Suite 2105  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10017  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5 1/4 inch diskette  
COMPUTER: PC/S LIMITED SYSTEM 200  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/604,965E  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/01270  
FILING DATE: June 27, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Wayne, Milton J.  
REGISTRATION NUMBER: 17,906  
REFERENCE/DOCKET NUMBER: U-WP-5212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-683-8150  
TELEFAX: 212-532-4285  
TELEX: 423794  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-604-965E-15

Query Match 48.1%; Score 25; DB 3; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EYGP 8  
|||  
Db 3 EYGP 6

RESULT 15  
5258287-43  
Patent No. 5258287  
APPLICANT: BAXTER, ROBERT C.; WOOD, WILLIAM I.  
TITLE OF INVENTION: DNA ENCODING AND METHODS OF PRODUCTION  
OF INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN BPS3  
NUMBER OF SEQUENCES: 58  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/171,623  
FILING DATE: 22-MAR-1988  
SEQ ID NO: 43:  
LENGTH: 7  
5258287-43

Query Match 48.1%; Score 25; DB 6; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EYGP 8  
|||  
Db 3 EYGP 6

Search completed: November 30, 2004, 08:15:59  
Job time : 28 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:41:32, Search time 11.6667 Seconds  
(without alignments)  
82.472 Million cell updates/sec

Title: US-10-008-377a-2  
Perfect score: 52  
Sequence: 1 GVLMWELFSL 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR.791.\*  
2: PIR1.\*  
3: PIR2.\*  
4: PIR3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	46.2	10	2	E49033 T-cell receptor ga
2	20	38.5	10	2	F49033 T-cell receptor ga
3	19	36.5	9	2	D57444 neurotrophin-3
4	18	34.6	9	2	D58350 gene C-mpl protein
5	17	32.7	6	2	PT0519 T-cell receptor be
6	17	32.7	7	2	PT0519 T-cell receptor be
7	17	32.7	10	2	PT0519 T-cell receptor be
8	17	32.7	10	2	PT0519 T-cell receptor be
9	16	30.8	5	2	PT0308 DNA polymerase - y
10	16	30.8	6	2	PT0308 DNA polymerase - y
11	16	30.8	6	2	PT0308 DNA polymerase - y
12	16	30.8	6	2	PT0308 DNA polymerase - y
13	16	30.8	6	2	PT0308 DNA polymerase - y
14	16	30.8	6	2	PT0308 DNA polymerase - y
15	16	30.8	6	2	PT0308 DNA polymerase - y
16	16	30.8	6	2	PT0308 DNA polymerase - y
17	16	30.8	6	2	PT0308 DNA polymerase - y
18	16	30.8	6	2	PT0308 DNA polymerase - y
19	16	30.8	6	2	PT0308 DNA polymerase - y
20	16	30.8	6	2	PT0308 DNA polymerase - y
21	16	30.8	6	2	PT0308 DNA polymerase - y
22	16	30.8	6	2	PT0308 DNA polymerase - y
23	16	30.8	6	2	PT0308 DNA polymerase - y
24	16	30.8	6	2	PT0308 DNA polymerase - y
25	16	30.8	6	2	PT0308 DNA polymerase - y
26	16	30.8	6	2	PT0308 DNA polymerase - y
27	16	30.8	6	2	PT0308 DNA polymerase - y
28	16	30.8	6	2	PT0308 DNA polymerase - y
29	16	30.8	6	2	PT0308 DNA polymerase - y

30	14	26.9	7	2	S33244 neuromodulatory pe
31	14	26.9	7	2	S33244 neuromodulatory pe
32	14	26.9	9	2	A61386 macrophage inhibi
33	14	26.9	10	2	A60624 angiotensin I - ja
34	14	26.9	10	2	A60624 T-cell receptor ga
35	14	26.9	10	2	A60624 T-cell receptor ga
36	14	26.9	10	2	A60624 T-cell receptor ga
37	14	26.9	10	2	A60624 T-cell receptor ga
38	14	26.9	10	2	A60624 T-cell receptor ga
39	14	26.9	10	2	A60624 T-cell receptor ga
40	14	26.9	10	2	A60624 T-cell receptor ga
41	14	26.9	10	2	A60624 T-cell receptor ga
42	14	26.9	10	2	A60624 T-cell receptor ga
43	14	26.9	10	2	A60624 T-cell receptor ga
44	14	26.9	10	2	A60624 T-cell receptor ga
45	14	26.9	10	2	A60624 T-cell receptor ga

## ALIGNMENTS

## RESULT 1

T-cell receptor gamma chain V-D-J region - human (fragment)  
C.Species: Homo sapiens (man)  
C.Date: 19-Dec-1993 #sequence\_revision 25-Aug-1995 #text\_change 21-Jul-2000  
C.Accession: E49033; D49033  
R.Morita, C.T.; Verma, S.; Aparicio, P.; Martinez, C.; Spits, H.; Brenner, M.B.  
Eur. J. Immunol. 21, 2999-3007, 1991  
A.Title: Functionally distinct subsets of human gamma/delta T cells.  
A.Reference number: A49033; PMID:92083926; PMID:1684157  
A.Accession: E49033  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-10 <MOR>  
A.Cross-references: GB:S72587; NID:G240696; PIDN:AA820630.1; PID:G240697  
A.Note: sequence extracted from NCBI backbone (NCBIN:72591, NCBIIP:72595)  
A.Accession: D49033  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-10 <MOR>  
A.Cross-references: GB:S72587; NID:G240696; PIDN:AA820630.1; PID:G240697  
A.Note: sequence extracted from NCBI backbone (NCBIN:72587, NCBIIP:72589)  
C.Keywords: T-cell receptor

## Query Match

Best local similarity 46.2%; Score 24; DB 2; Length 10;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 LMEWELFSL 10  
DB 3 LMEWELFSL 9

## RESULT 2

T-cell receptor gamma chain V-D-J region - human (fragment)  
C.Species: Homo sapiens (man)  
C.Date: 19-Dec-1993 #sequence\_revision 17-Mar-2000 #text\_change 17-Mar-2000  
C.Accession: F49033  
R.Morita, C.T.; Verma, S.; Aparicio, P.; Martinez, C.; Spits, H.; Brenner, M.B.  
Eur. J. Immunol. 21, 2999-3007, 1991  
A.Title: Functionally distinct subsets of human gamma/delta T cells.  
A.Reference number: A49033; PMID:92083926; PMID:1684157  
A.Accession: F49033  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-10 <MOR>  
A.Cross-references: GB:S72605; NID:G240700; PIDN:AA820632.1; PID:G240701  
A.Note: sequence extracted from NCBI backbone (NCBIN:72605, NCBIIP:72606)  
C.Keywords: T-cell receptor

## Query Match

38.5%; Score 20; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LME 6  
|||  
Db 3 LME 5

## RESULT 3

D57444  
neuropeptide Grp-Asp B4 - two-spotted cricket

C/Species: Gryllus bimaculatus (two-spotted cricket)  
C/Date: 26-Jun-1996 #sequence\_revision 26-Jan-1996 #ext\_change 09-Jul-2004

C/Accession: D57444  
R/Lorenz, M.W.; Kellner, R.; Hoffmann, K.H.

J. Biol. Chem. 270: 21103-21108, 1995

A/Title: A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cricket

A/Reference number: A57444; MUID:95403341; PMID:7673141

A/Accession: D57444

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-9 <LOR>

A/Cross-references: UNIPROT:Q7M3N6

## Query Match

Best Local Similarity 36.5%; Score 19; DB 2; Length 9;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 WEIF 8  
|||  
Db 2 WEIF 5

## RESULT 4

I58350  
gene c-mpl protein - mouse (fragment)

C/Species: Mus sp. (mouse)  
C/Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #ext\_change 28-Feb-1997

C/Accession: I58350  
R/Alexander, W.S.; Dunn, A.R.

Oncogene 10, 795-803, 1995

A/Title: Structure and transcription of the genomic locus encoding murine c-mpl, a receptor

A/Reference number: I58350; MUID:95166571; PMID:7862460

A/Accession: I58350

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-9 <RMS>

A/Cross-references: GB:S76841; NID:912992

C/Accession: I58350

A/Status: preliminary

A/Residues: 1-9 <RMS>

A/Cross-references: GB:S76841; NID:912992

C/Accession: I58350

A/Status: preliminary

A/Residues: 1-9 <RMS>

A/Cross-references: GB:S76841; NID:912992

A/Residues: 1-6 <PEZ>  
A/Experimental source: adult thymus, strain BALB/c  
C/Keywords: T-cell receptor

Query Match  
Best Local Similarity 32.7%; Score 17; DB 2; Length 6;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 LME 6  
|||  
Db 4 LMD 6

## RESULT 6

PX0008  
glucuronosyltransferase (EC 2.4.1.17), hepatic - rat (fragment)

A/Alternate names: UDP-glucuronosyltransferase

C/Species: Rattus norvegicus (Norway rat)  
C/Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #ext\_change 07-Feb-1997

C/Accession: PX0008  
R/Yokota, H.; Yuasa, A.; Sato, R.

J. Biochem. 104, 531-536, 1988

A/Title: Purification and properties of a form of UDP-glucuronosyltransferase from liver

A/Reference number: PX0008; MUID:89197852; PMID:3149280

A/Accession: PX0008

A/Molecule type: protein

A/Residues: 1-7 <FOX>

A/Cross-references: glycosyltransferase, liver

## Query Match

Best Local Similarity 32.7%; Score 17; DB 2; Length 7;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLM 5  
|||  
Db 3 LVM 6

## RESULT 7

S71868  
N/A alternate names: glutathione S-transferase class mu 4

C/Species: Sus scrofa domestica (domestic pig)  
C/Date: 19-Mar-1998 #sequence\_revision 13-Sep-1998 #ext\_change 09-Jul-2004

C/Accession: S71868  
R/Roum, P.; Anglade, P.; Depraetere, L.; Tulliez, J.

Biochem. J. 317, 879-884, 1998

A/Title: Characterization of pig liver glutathione S-transferases using HPLC-electrospr

A/Reference number: S71868; MUID:96332484; PMID:9760377

A/Accession: S71868

A/Molecule type: protein

A/Residues: 1-10 <ROU>

A/Cross-references: UNIPROT:Q7M3B8

C/Comment: At least five species-independent classes of cytosolic glutathione transferases

C/Complex: dimer

C/Function: catalyzes the nucleophilic conjugation of intracellular glutathione to a

A/Pathway: detoxification; xenobiotics metabolism

A/Note: increased hydrophilicity of GSH-conjugates facilitates their further metabolism

C/Suprafamily: glutathione transferase

C/Keywords: dimer; transferase

Query Match  
Best Local Similarity 32.7%; Score 17; DB 2; Length 10;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEI 7  
|||  
Db 6 WDI 8

RESULT 8  
 S63696  
 DNA polymerase - yeast (*Kluyveromyces marxianus* var. *lactis*) killer plasmid pgKL2 (fragment)  
 C:Species: *Kluyveromyces marxianus* var. *lactis*, *Candida sp.*  
 C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 25-Apr-1997  
 C:Accession: S63696  
 R:Trivedi, M.; Hiraishi, H.; Takesako, T.; Tanase, S.; Gunge, N.  
 Year: 12, 241-246, 1996  
 A:Title: The terminal protein of the linear DNA plasmid pgKL2 shares an N-terminal domain  
 A:Reference number: S63696; PMID:97060015; PMID:8904336  
 A:Accession: S63696  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-10 <TAX>  
 A>Note: the source is designated as *Kluyveromyces lactis*

Query Match  
 Best Local Similarity 32.7%; Score 17; DB 2; Length 10;  
 Best Local 60.0%; Pred. No. 4.2e+03;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 VLLME 6  
 DB 1 VXXME 5

RESULT 9  
 PT0308  
 Ig heavy chain CRD3 region (clone 6-88) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0308  
 R:Yanada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Catton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
 A:Reference number: PT0222; PMID:91108337; PMID:1899102  
 A:Accession: PT0308  
 A:Molecule type: DNA  
 A:Residues: 1-5 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotrimer; immunoglobulin

Query Match  
 Best Local Similarity 30.8%; Score 16; DB 2; Length 5;  
 Best Local 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WE 6  
 DB 2 WE 3

RESULT 10  
 S66195  
 alcohol dehydrogenase (EC 1.1.1.1) class III low affinity form - cod (Gadus sp.) (fragment)  
 C:Species: *Gadus* sp. (cod)  
 C:Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 12-Jun-1998  
 C:Accession: S66195  
 R:Hilmy, L.; Hackett, M.; Safgat, J.; Danielsson, O.; Iida, J.; Hendrickson, R.C.;  
 FEBS Lett. 367, 237-240, 1995  
 A:Title: Multiplicity of N-terminal structures of medium-chain alcohol dehydrogenases. N  
 zyme.  
 A:Reference number: S66191; PMID:95331382; PMID:7607314  
 A:Accession: S66195  
 A:Molecule type: protein  
 A:Residues: 1-6 <HJB>  
 C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology  
 C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match  
 Best Local Similarity 30.8%; Score 16; DB 2; Length 6;  
 Best Local 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WE 6  
 DB 1 WE 6

DB 5 WE 6  
 RESULT 11  
 B34835  
 dhaA protein - *Pseudomonas aeruginosa* (fragment)  
 C:Species: *Pseudomonas aeruginosa*  
 C:Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 08-Oct-1999  
 C:Accession: B34835  
 R:Yee, T.W.; Smith, D.W.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 1278-1282, 1990  
 A:Title: Pseudomonas chromosomal replication origins: a bacterial class distinct from E  
 A:Reference number: A34835; PMID:90160310; PMID:2106132  
 A:Accession: B34835  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-6 <YEE>  
 A:Cross-references: GB:M30125; NID:9151419; PID:AAA25916.1; PID:9151421  
 C:Keywords: DNA binding

Query Match  
 Best Local Similarity 30.8%; Score 16; DB 2; Length 6;  
 Best Local 75.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLLM 5  
 DB 3 VELL 6

RESULT 12  
 A11263  
 dihydrofolate reductase (EC 1.5.1.3) / thymidylate synthase (EC 2.1.1.45) - *Plasmodium falciparum*  
 C:Species: *Plasmodium falciparum*  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Mar-1996  
 C:Accession: A11263  
 R:Peterson, D.S.; Walliker, D.; Wellens, T.E.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 9114-9118, 1988  
 A:Title: Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase  
 A:Reference number: A94217; PMID:89057886; PMID:2904149  
 A:Accession: A11263  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-6 <PET>  
 C:Keywords: methyltransferase; NADP; oxidoreductase

Query Match  
 Best Local Similarity 30.8%; Score 16; DB 2; Length 6;  
 Best Local 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WE 6  
 DB 3 WE 4

RESULT 13  
 B31263  
 dihydrofolate reductase (EC 1.5.1.3) / thymidylate synthase (EC 2.1.1.45) - *Plasmodium falciparum*  
 C:Species: *Plasmodium falciparum*  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Mar-1996  
 C:Accession: B31263  
 R:Peterson, D.S.; Walliker, D.; Wellens, T.E.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 9114-9118, 1988  
 A:Title: Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase  
 A:Reference number: A94217; PMID:89057886; PMID:2904149  
 A:Accession: B31263  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-6 <PET>  
 C:Keywords: methyltransferase; NADP; oxidoreductase

Query Match  
 Best Local Similarity 30.8%; Score 16; DB 2; Length 6;  
 Best Local 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WE 6  
||  
Db 3 WE 4

## RESULT 14

B35640  
cerebellar degeneration-related protein - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 28-Sep-1990 #sequence\_revision 28-Sep-1990 #text\_change 24-Jun-1993  
C/Accession: B35640  
R:Chen, Y.T.; Rettig, W.J.; Yemamandra, A.K.; Kozak, C.A.; Chaganti, R.S.K.; Posner, J.E  
Proc. Natl. Acad. Sci. U.S.A. 87, 3077-3081, 1990  
A>Title: Cerebellar degeneration-related antigen: a highly conserved neuroectodermal marker  
A/Reference number: A35640; MUID:50222173; PMID:2326268  
A/Accession: B35640  
A/Status: Preliminary  
A/Molecule type: mRNA  
A/Residues: 1-6 <CHE>

Query Match 30.8%; Score 16; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Gaps 0;

QY 5 WE 6  
||  
Db 2 WE 3

## RESULT 15

S09652  
hypothetical protein (aacC2 3' region) - Enterobacter cloacae (fragment)  
C:Species: Enterobacter cloacae  
C>Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 08-Oct-1999  
C/Accession: S09652  
R:Wiegant, J.S.; Kerebelar-van Gaalen, P.A.G.; van de Klundert, J.A.M.  
Antimicrob. Agents Chemother. 33, 1153-1159, 1989  
A>Title: Nucleotide sequence of the aacC2 gene, a gentamicin resistance determinant invc  
A/Reference number: S09651; MUID:50024972; PMID:2552900  
A/Accession: S09652  
A/Molecule type: DNA  
A/Residues: 1-7 <VLI>  
A/Cross-references: EMBL:X51534; NID:940878; PID:CAA35914.1; PID:9581034

Query Match 30.8%; Score 16; DB 2; Length 7;  
Best Local Similarity 25.0%; Pred. No. 2.8e+05;  
Matches 1; Conservative 3; Mismatches 0; Gaps 0;

QY 2 VLLW 5  
:::  
Db 1 MLW 4

Search completed: November 30, 2004, 08:14:26  
Job time: 12.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:40:12 / Search time 78.3333 Seconds  
(without alignments)  
73.452 Million cell updates/sec

Title: US-10-008-377A-2  
Perfect score: 52  
Sequence: 1 GVLWEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1825161 seqs, 57537466 residues

Total number of hits satisfying chosen parameters: 2971

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: UniProt\_02: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44.2	10	2	06JL97	06JL97	Neisseria g
2	44.2	10	2	AA516521	AA516521	Neisseria g
3	42.3	10	2	Q8JFE7	Q8JFE7	flicella al
4	42.3	10	2	Q8JFE7	Q8JFE7	flicella al
5	38.5	10	2	Q8JFE7	Q8JFE7	flicella al
6	38.5	10	2	Q8JFE7	Q8JFE7	flicella al
7	36.5	10	2	Q8JFE7	Q8JFE7	flicella al
8	36.5	10	2	Q8JFE7	Q8JFE7	flicella al
9	36.5	10	2	Q8JFE7	Q8JFE7	flicella al
10	36.5	10	2	Q8JFE7	Q8JFE7	flicella al
11	34.6	10	1	AAE_CARGI	AAE_CARGI	mus sapien
12	34.6	10	1	AAE_CARGI	AAE_CARGI	mus sapien
13	32.7	10	2	Q8JFE7	Q8JFE7	flicella al
14	32.7	10	2	Q8JFE7	Q8JFE7	flicella al
15	32.7	10	2	Q8JFE7	Q8JFE7	flicella al
16	32.7	10	2	Q8JFE7	Q8JFE7	flicella al
17	32.7	10	2	Q8JFE7	Q8JFE7	flicella al
18	32.7	10	2	Q8JFE7	Q8JFE7	flicella al
19	32.7	10	2	Q8JFE7	Q8JFE7	flicella al
20	32.7	10	2	Q8JFE7	Q8JFE7	flicella al
21	32.7	10	2	Q8JFE7	Q8JFE7	flicella al
22	32.7	10	2	Q8JFE7	Q8JFE7	flicella al
23	32.7	10	2	Q8JFE7	Q8JFE7	flicella al
24	30.8	10	1	AAE_CARGI	AAE_CARGI	mus sapien
25	30.8	10	1	AAE_CARGI	AAE_CARGI	mus sapien
26	30.8	10	1	AAE_CARGI	AAE_CARGI	mus sapien
27	30.8	10	1	AAE_CARGI	AAE_CARGI	mus sapien
28	30.8	10	1	AAE_CARGI	AAE_CARGI	mus sapien
29	30.8	10	1	AAE_CARGI	AAE_CARGI	mus sapien
30	30.8	10	1	AAE_CARGI	AAE_CARGI	mus sapien
31	30.8	10	1	AAE_CARGI	AAE_CARGI	mus sapien

## ALIGNMENTS

32	15	30.8	9	2	Q70YA3	Q70YA3	aeollanthus
33	16	30.8	9	2	Q94YG2	Q94YG2	varanus ind
34	16	30.8	9	2	Q85723	Q85723	simian sarc
35	16	30.8	9	2	CAD45451	CAD45451	aeollanth
36	16	30.8	10	2	Q7SA62	Q7SA62	neurospora
37	16	30.8	10	2	Q8SHB4	Q8SHB4	furcifer ve
38	16	30.8	10	2	Q8SHB7	Q8SHB7	furcifer ou
39	16	30.8	10	2	Q8SHC0	Q8SHC0	furcifer la
40	16	30.8	10	2	Q8SHC3	Q8SHC3	notophthalm
41	15.5	29.8	10	2	Q9BAX0	Q9BAX0	chamaeleo
42	15.5	29.8	10	2	Q8SHD2	Q8SHD2	chamaeleo
43	15.5	29.8	10	2	Q8SHD5	Q8SHD5	chamaeleo
44	15.5	29.8	10	2	Q8SHD8	Q8SHD8	chamaeleo
45	15.5	29.8	10	2	Q8SHH4	Q8SHH4	chamaeleo

  

RESULT 1	Q6JL97	PREDIMINARY;	PRT;	10 AA.
ID	Q6JL97;			
AC	Q6JL97;			
DT	05-JUL-2004 (TREMBLrel. 27, Created)			
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)			
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)			
DE	Nucl (Fragment).			
GN	Name=Nucl.			
OS	Neisseria gonorrhoeae.			
OC	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;			
OC	Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=485;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MS11;			
RX	PubMed=15084227;			
RA	Snyder L.A., Davies J.K., Saunders N.J.;			
RT	"Microarray genotyping of key experimental strains of Neisseria			
RT	gonorrhoeae reveals gene complement diversity and five new neisserial			
RT	genes associated with Minimal Mobile Elements."			
RL	BMC Genomics 5:23-23(2004).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MS11;			
RA	Snyder L.A.S., Davies J.K., Saunders N.J.;			
RT	Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY386266; AA516521.1; -			
FT	NON_TER			
SQ	SEQUENCE 10 AA; 1227 MW; BACCB28379D1A6 CRC64;			

  

Query Match	44.2%; Score 23; DB 2; Length 10;
Best Local Similarity	37.5%; Pred. No. 2.6e+03;
Matches	3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

  

QY	1 GVLWEIF 8
DB	2 GMTFWGLF 9

  

RESULT 2	AA516521	PREDIMINARY;	PRT;	10 AA.
ID	AA516521;			
AC	AA516521;			
DT	20-MAY-2004 (TREMBLrel. 27, Created)			
DT	20-MAY-2004 (TREMBLrel. 27, Last sequence update)			
DT	20-MAY-2004 (TREMBLrel. 27, Last annotation update)			
DE	Nucl (Fragment).			
GN	Nucl.			
OS	Neisseria gonorrhoeae.			
OC	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;			
OC	Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=485;			
RN	[1]			

```

RP SEQUENCE FROM N.A.
RC STRAIN=MS11;
RX PubMed=15084227;
RA Snyder L.A., Davies J.K., Saunders N.J.;
RT "Microarray genotyping of key experimental strains of Neisseria
RT gonorrhoeae reveals gene complement diversity and five new neisseria
RT genes associated with Minimal Mobile Elements.";
RL BMC Genomics 5:23-23(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MS11;
RA Snyder L.A.S., Davies J.K., Saunders N.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY386266; AAS16521.1; -.
FT NON_TER
SQ SEQUENCE 10 AA; 1227 MW; BACCB286379D1A6 CRC64;

Query Match
Best Local Similarity 44.2%; Score 23; DB 2; Length 10;
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVLWLF 8
DB 2 GMTFWGLF 9

RESULT 3
ID Q8UFE7 PRELIMINARY; PRT; 10 AA.
AC Q8UFE7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 05-JUN-2004 (TREMBlrel. 22, Last annotation update)
DE Myelin proteolipid protein (Fragment).
OS Piceidula albicollis (Collared flycatcher).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Passeriformes; Muscicapidae; Ficedula.
OC NCBI_TaxID=59894;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21918460; PubMed=11918793;
RA Primer C.R., Borge T., Lindell J., Saetre G.P.;
RT "Single-nucleotide polymorphism characterization in species with
RT limited available sequence information: high nucleotide diversity
RT revealed in the avian genome.";
RL Mol. Ecol. 11:603-612(2002).
DR EMBL; AF454217; AAM22903.1; -.
DR EMBL; AF454218; AAM22904.1; -.
FT NON_TER
SQ SEQUENCE 10 AA; 1206 MW; 1FAAC9676732C86B CRC64;

Query Match
Best Local Similarity 42.3%; Score 22; DB 2; Length 10;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVLW 5
DB 4 GVLW 8

RESULT 4
ID Q8UJ33 PRELIMINARY; PRT; 10 AA.
AC Q8UJ33;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Myelin proteolipid protein (Fragment).
OS Ficedula hypoleuca.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Passeriformes; Muscicapidae; Ficedula.
OC NCBI_TaxID=46689;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21918460; PubMed=11918793;
RA Primer C.R., Borge T., Lindell J., Saetre G.P.;
RT "Single-nucleotide polymorphism characterization in species with
RT limited available sequence information: high nucleotide diversity
RT revealed in the avian genome.";
RL Mol. Ecol. 11:603-612(2002).
DR EMBL; AF454216; AAM22902.1; -.
FT NON_TER
SQ SEQUENCE 10 AA; 1206 MW; 1FAAC9676732C86B CRC64;

Query Match
Best Local Similarity 42.3%; Score 22; DB 2; Length 10;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVLW 5
DB 4 GVLW 8

RESULT 5
ID Q85406 PRELIMINARY; PRT; 8 AA.
AC Q85406;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Coxsiella burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxsiellaceae; Coxsiella.
OC NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile Phase I;
RX MEDLINE=98348442; PubMed=9633477;
RA Williams H., Jaeger C., Baljer G.;
RT "Physical and genetic map of the obligate intracellular bacterium
RT Coxsiella burnetii.";
RL J. Bacteriol. 180:3816-3822(1998).
DR EMBL; AF064963; AAD09947.1; -.
KM Hypothetical protein.
FT NON_TER
SQ SEQUENCE 8 AA; 993 MW; 046BSAA53772727 CRC64;

Query Match
Best Local Similarity 38.5%; Score 20; DB 2; Length 8;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLLW 5
DB 1 LLLW 4

RESULT 6
ID Q8SHF6 PRELIMINARY; PRT; 10 AA.
AC Q8SHF6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN Name=COI;
OS Chamaeleo melleri (Meller's chameleon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae; Chamaeleo.
OC NCBI_TaxID=179915;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22169767; PubMed=12162400;

```



RA Townsend T., Larson A.;  
 RT "Molecular phylogenetics and mitochondrial genomic evolution in the  
 RT chameleons (Reptilia, Squamata).";  
 RL Mol. Phylogenet. Evol. 23:22-36(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Townsend T.M., Larson A.L.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF448755; AAL90547.1; -  
 DR GO: GO:0005739; Mitochondrion; IEA.  
 KM Mitochondrion.  
 FT NON TER  
 SQ SEQUENCE 10 AA; 1276 MW; 58218E2733772727 CRC64;

Query Match 38.5%; Score 20; DB 2; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 9.1e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLW 5  
 : : :  
 Db 3 LILW 6

RESULT 7  
 Q8JEB1 PRELIMINARY; PRT; 7 AA.  
 AC Q8JEB1;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Truncated pol protein (Fragment).  
 GN Name-pol;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 ON NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22056123; PubMed=12060770;  
 RA Beerenwinkel N., Schmidt B., Walter H., Kaiser R., Lengauer T.,  
 RA Hoffmann D., Korn K., Selbig J.;  
 RT "Diversity and complexity of HIV-1 drug resistance: a bioinformatics  
 RT approach to predicting phenotype from genotype.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:8271-8276(2002).  
 DR EMBL: AF347267; AAK32344.1; -  
 FT NON TER  
 SQ SEQUENCE 7 AA; 885 MW; 76C37731A046C700 CRC64;

Query Match 36.5%; Score 19; DB 2; Length 7;  
 Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLW 6  
 : : :  
 Db 3 LILW 7

RESULT 8  
 Q7M3N6 PRELIMINARY; PRT; 9 AA.  
 ID Q7M3N6;  
 AC Q7M3N6;  
 DT 01-MAR-2004 (TREMBlrel. 26, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Neuropeptide Grb-AST B4.  
 GN Neuropeptide Grb-AST B4.  
 OS Gyllius bimaculatus (Two-spotted cricket).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygotia;  
 OC Neoptera; Orthoptera; Orthoptera; Ensifera; Gryllidae; Gryllinae;  
 OC Gyllius.  
 ON NCBI\_TaxID=6999;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95403341; PubMed=7673141;  
 RA Lorenz M.W., Kellner R., Hoffmann K.H.;

RT "A family of neuropeptides that inhibit juvenile hormone biosynthesis  
 RT in the cricket, Gyllius bimaculatus.";  
 RL J. Biol. Chem. 270:21103-21108(1995).  
 DR PIR: D57444; D57444.1;  
 SQ SEQUENCE 9 AA; 1175 MW; 3860871E9D40B03 CRC64;

Query Match 36.5%; Score 19; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 WEIF 8  
 : : :  
 Db 2 WEIF 5

RESULT 9  
 Q8WTT4 PRELIMINARY; PRT; 10 AA.  
 ID Q8WTT4;  
 AC Q8WTT4;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Dyscrophin (Fragment).  
 GN Name=dyscrophin; Synonyms=DMD;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ditea S.D., Klamut H.T., Ray P.N., Worton R.G.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Beggs A.H., Koenig M., Kunkel L.M.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA den Dunnen J.T.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF276053; AAL35752.1; -  
 DR EMBL: AF213401; AAL61549.1; -  
 FT NON TER  
 SQ SEQUENCE 10 AA; 1340 MW; 2B999202CB1B0363 CRC64;

Query Match 36.5%; Score 19; DB 2; Length 10;  
 Best Local Similarity 60.0%; Pred. No. 1.4e+04;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLW 6  
 : : :  
 Db 1 MLWWE 5

RESULT 10  
 Q8OX07 PRELIMINARY; PRT; 9 AA.  
 ID Q8OX07;  
 AC Q8OX07;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Thrombopoietin receptor (Fragment).  
 GN Name=c-mpl;  
 OS Mus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10095;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95165571; PubMed=7862460;  
 RA Alexander W.S., Dunn A.R.;  
 RT "Structure and transcription of the genomic locus encoding murine c-  
 RT Mpl, a receptor for thrombopoietin.";

RL Oncogene 10:795-803(1995).  
 DR EMBL: S76841; A03198.1; -  
 DR GO: GO:0004872; F:receptor activity; IEA.  
 RT Receptor.  
 RT NON\_TER.  
 SQ SEQUENCE 9 AA; 1081 MW; 2DPF99C72DC355A7 CRC64;

Query Match 34.6%; Score 18; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 WEIR 8  
 DB 4 MALV 7

RESULT 11  
 APE\_CAPI STANDARD; PRT; 10 AA.

AC P80474;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE Aminopeptidase (EC 3.4.11.-) (Fragment).  
 OS Capnocytophaga gingivalis.  
 OC Bacteria; Bacteroidetes; Flavobacteriales;  
 OC Flavobacteriaceae; Capnocytophaga.  
 OX NCBI\_TaxID=1017;  
 RN [1]  
 RP SEQUENCE.

RC STRAIN=ATCC 33624;  
 RX MEDLINE=96118234; PubMed=8574402;  
 RA Spratt D.A., Greenman J., Schaffer A.G.;  
 RT Capnocytophaga gingivalis aminopeptidase: a potential virulence factor.  
 RL Microbiology 141:3087-3093(1995).  
 CC -1- FUNCTION: Aminopeptidase which hydrolyzes substrates with free N-terminal amino acid residues but not N-terminal blocked ones.  
 CC Optimum activity is measured at pH 7.5. May be important in the nutrition and pathogenesis of the organism in the human oral cavity.  
 CC -1- Cofactor: Requires magnesium or calcium.

CC Aminopeptidase; Calcium; Direct protein sequencing; Hydrolase;  
 KM Magnesium.  
 FT NON\_TER 1 1  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1306 MW; 00C0A6DB43772694 CRC64;

Query Match 34.6%; Score 18; DB 1; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 2.1e+04;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LMWEI 7  
 DB 4 MLWYV 8

RESULT 12  
 Q9ESUS PRELIMINARY; PRT; 10 AA.

AC Q9ESUS;  
 DT 01-MAR-2001 (TRENBLREL 16, Created)  
 DT 01-MAR-2001 (TRENBLREL 16, Last sequence update)  
 DT 01-JUN-2003 (TRENBLREL 24, Last annotation update)  
 DE Fas death receptor (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Soturognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/He;  
 RX MEDLINE=20127858; PubMed=10660538;

RA Munsch D., Watanabe-Fukunaga R., Bourdon J.C., Nagata S., May E.,  
 RA Yoshish-Rouach E., Reisdorf P.;  
 RT "Human and mouse Fas (Apo-1/CD95) death receptor genes each contain a p53-responsive element that is activated by p53 mutants unable to induce apoptosis."  
 RT J. Biol. Chem. 275:3867-3872(2000).  
 RL EMBL: AF282865; A002410.1; -  
 DR GO: GO:0004872; F:receptor activity; IEA.  
 RT Receptor.  
 RT NON\_TER.  
 SQ SEQUENCE 10 AA; 1242 MW; 22145532CDG37043 CRC64;

Query Match 34.6%; Score 18; DB 2; Length 10;  
 Best Local Similarity 28.6%; Pred. No. 2.1e+04;  
 Matches 2; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 LMWEI 10  
 DB 4 IMAVLPL 10

RESULT 13  
 O81VK3 PRELIMINARY; PRT; 8 AA.

AC O81VK3;  
 DT 01-MAR-2003 (TRENBLREL 23, Created)  
 DT 01-MAR-2003 (TRENBLREL 23, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL 23, Last annotation update)  
 DE Steerin2 (Fragment).  
 GN Name=STEERIN2;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Peeters P.J., Verhasselt P., Moenchars D.W., Luyten W.H.M.L.,  
 RA Geysen J.J.G.H.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ488208; CAD32561.1; -  
 FT NON\_TER 8 8  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 1030 MW; E8A7245417737726 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LMW 5  
 DB 1 MLW 3

RESULT 14  
 Q38366 PRELIMINARY; PRT; 9 AA.

AC Q38366;  
 DT 01-NOV-1996 (TRENBLREL 01, Created)  
 DT 01-NOV-1996 (TRENBLREL 01, Last sequence update)  
 DT 01-DEC-2001 (TRENBLREL 19, Last annotation update)  
 DE E gene product (Fragment).  
 OS Bacteriophage phi-X174.  
 OC Viruses; ssDNA viruses; Microviridae; Microvirus.  
 OX NCBI\_TaxID=10847;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89118956; PubMed=2963134;  
 RA Buckley K.J., Hayaishi M.;  
 RT "Role of premature translational termination in the regulation of RT expression of the phi X174 lysis gene."  
 RT J. Mol. Biol. 198:599-607(1987).  
 RL EMBL: X07809; CA30668.1; -  
 FT NON\_TER 9 9  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 1207 MW; C093B37731B36412 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0;

QY 4 LME 6  
 11:  
 Db 6 LMD 8

## RESULT 15

Q25355 PRELIMINARY; PRT; 10 AA.  
 AC Q25355;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE Viteillogenin A (Fragment).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridoidea;  
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.  
 OX NCBI\_Taxid=7004;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86003971; PubMed=2820677;  
 RA Locke J., White B.N., Wyatt G.R.;  
 RT "Cloning and 5' end nucleotide sequences of two juvenile hormone-  
 inducible vitellogenin genes of the African migratory locust.",  
 RL DNA 6:331-342(1987).  
 DR EMBL; M17333; AAA29284.1; -.  
 FT NON TER 10  
 SQ SEQUENCE 10 AA; 1128 MW; D1B31177272042CD CRC64;

Query Match 32.7%; Score 17; DB 2; Length 10;  
 Best Local Similarity 28.6%; Pred. No. 3.2e+04;  
 Matches 2; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 LMEFSL 10  
 11:  
 Db 1 MMAVILL 7

Search completed: November 30, 2004, 08:13:40  
 Job time : 79.333 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:41:57 ; Search time 27 seconds  
(without alignments)  
24.562 Million cell updates/sec

Title: US-10-008-377A-2

Perfect score: 52

Sequence: 1 GVLMWTFSL 10

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 110780

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	59.6	9	4	US-09-462-453-22
2	53.8	9	4	US-09-667-857-438
3	53.8	10	4	US-09-667-857-421
4	53.8	10	4	US-09-667-857-433
5	51.9	8	4	US-09-786-4808-42
6	50.0	8	3	US-09-082-2798-1488
7	50.0	8	3	US-09-315-3048-1642
8	50.0	8	4	US-09-834-784-1488
9	50.0	8	4	US-09-350-441C-1643
10	48.1	7	3	US-08-335-8657-18
11	46.2	6	4	US-09-121-211-12
12	46.2	8	2	US-08-357-642A-4
13	46.2	8	2	US-08-357-642A-14
14	46.2	8	2	US-08-460-626-4
15	46.2	8	2	US-08-460-626-14
16	46.2	8	2	US-08-460-626-18
17	46.2	8	3	US-08-836-325-4
18	46.2	8	4	US-09-457-571-4
19	46.2	9	3	US-08-159-339A-648
20	46.2	9	4	US-09-682-325-22
21	46.2	9	4	US-09-682-325-24
22	46.2	10	2	US-08-751-300-28
23	46.2	10	2	US-08-751-300-29
24	46.2	10	3	US-08-159-339A-727
25	46.2	10	4	US-08-728-742A-61
26	46.2	10	4	US-09-043-813-28
27	46.2	10	4	US-09-043-813-29

28	23	44.2	6	1	US-07-802-667-27	Sequence 27, Appl
29	23	44.2	6	1	US-07-802-667-28	Sequence 28, Appl
30	23	44.2	6	4	US-09-282-029A-32	Sequence 32, Appl
31	23	44.2	6	4	US-09-282-029A-108	Sequence 108, Appl
32	23	44.2	6	4	US-09-185-908-32	Sequence 32, Appl
33	23	44.2	6	4	US-09-185-908-108	Sequence 108, Appl
34	23	44.2	7	4	US-09-282-029A-33	Sequence 33, Appl
35	23	44.2	7	4	US-09-282-029A-109	Sequence 109, Appl
36	23	44.2	7	4	US-09-185-908-33	Sequence 33, Appl
37	23	44.2	7	4	US-09-185-908-109	Sequence 109, Appl
38	23	44.2	8	2	US-08-641-314C-7	Sequence 7, Appl
39	23	44.2	8	3	US-09-082-2798-1433	Sequence 1433, Ap
40	23	44.2	8	3	US-09-082-2798-1474	Sequence 1474, Ap
41	23	44.2	8	3	US-09-082-2798-1484	Sequence 1484, Ap
42	23	44.2	8	3	US-09-082-2798-1485	Sequence 1485, Ap
43	23	44.2	8	3	US-09-082-2798-1487	Sequence 1487, Ap
44	23	44.2	8	3	US-09-315-3048-1433	Sequence 1433, Ap
45	23	44.2	8	3	US-09-315-3048-1562	Sequence 1562, Ap

# ALIGNMENTS

RESULT 1  
US-09-462-453-22  
Sequence 22, Application US/09462453  
Patent No. 6723695  
GENERAL INFORMATION:  
APPLICANT: BUREAUX, SCOTT RENTON  
APPLICANT: KHANNA, RAJIV  
APPLICANT: SHERITT, MARTINA ALISON  
TITLE OF INVENTION: CTU PEPTIDES FROM EBV  
FILE REFERENCE: FBRC:010  
CURRENT APPLICATION NUMBER: US/09/462,453  
CURRENT FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 22  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Epstein Barr Virus  
US-09-462-453-22

Query Match 59.6%; Score 31; DB 4; Length 9;  
Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VILMEI 7  
Db 4 VILMEV 9

RESULT 2  
US-09-667-857-438  
Sequence 438, Application US/09667857  
Patent No. 6699664  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
APPLICANT: Aligate, Paul A.  
APPLICANT: Fling, Steven P.  
APPLICANT: Retter, Marc W.  
APPLICANT: Ranger, Gary Richard  
APPLICANT: Reed, Steven G.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darriek  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.462C5  
CURRENT APPLICATION NUMBER: US/09/667,857  
CURRENT FILING DATE: 2000-09-20  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 438  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-667-857-438

Query Match  
Best Local Similarity 53.8%; Score 28; DB 4; Length 9;  
Best Local Similarity 44.4%; Pred. No. 3.8e+05;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLMEIFSL 10  
: |||||:  
Db 1 ILFWSIISI 9

RESULT 3  
US-09-667-857-421  
Sequence 421, Application US/09667857  
Patent No. 6699664  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
APPLICANT: Algate, Paul A.  
APPLICANT: Fling, Steven P.  
APPLICANT: Retter, Marc W.  
APPLICANT: Ranger, Gary Richard  
APPLICANT: Reed, Steven G.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darick  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.462CS  
CURRENT APPLICATION NUMBER: US/09/667,857  
CURRENT FILING DATE: 2000-09-20  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 421  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-667-857-421

Query Match  
Best Local Similarity 53.8%; Score 28; DB 4; Length 10;  
Best Local Similarity 44.4%; Pred. No. 63;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLMEIFSL 10  
: |||||:  
Db 1 ILFWSIISI 9

RESULT 4  
US-09-667-857-433  
Sequence 433, Application US/09667857  
Patent No. 6699664  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
APPLICANT: Algate, Paul A.  
APPLICANT: Fling, Steven P.  
APPLICANT: Retter, Marc W.  
APPLICANT: Ranger, Gary Richard  
APPLICANT: Reed, Steven G.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darick  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.462CS  
CURRENT APPLICATION NUMBER: US/09/667,857  
CURRENT FILING DATE: 2000-09-20  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 433

LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-667-857-433

Query Match  
Best Local Similarity 53.8%; Score 28; DB 4; Length 10;  
Best Local Similarity 44.4%; Pred. No. 63;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLMEIFSL 10  
: |||||:  
Db 2 ILFWSIISI 10

RESULT 5  
US-09-786-480B-42  
Sequence 42, Application US/09786480B  
Patent No. 6730825  
GENERAL INFORMATION:  
APPLICANT: Goldsbrough, Andrew  
APPLICANT: Colliver, Steve  
TITLE OF INVENTION: Isoforms of Starch Branching Enzyme II (SBE-IIA and SBE-IIB) From  
FILE REFERENCE: 11951.0005.PCUS00 MSB:005  
CURRENT APPLICATION NUMBER: US/09/786,480B  
CURRENT FILING DATE: 2002-05-25  
PRIOR APPLICATION NUMBER: PCT/GB99/03011  
PRIOR FILING DATE: 1999-09-09  
PRIOR APPLICATION NUMBER: EP 98307337.0  
PRIOR FILING DATE: 1998-09-10  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 42  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Triticum aestivum  
US-09-786-480B-42

Query Match  
Best Local Similarity 51.9%; Score 27; DB 4; Length 8;  
Best Local Similarity 80.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 LMEIF 8  
: |||||:  
Db 2 VMEIF 6

RESULT 6  
US-09-082-279B-1488  
Sequence 1488, Application US/09082279B  
Patent No. 6258782  
GENERAL INFORMATION:  
APPLICANT: Barney, Shawn  
APPLICANT: Guthrie, Kelly  
APPLICANT: Merutka, Gene  
APPLICANT: Anwer, Mohamed  
APPLICANT: Lambert, Dennis  
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED  
TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES  
FILE REFERENCE: 7872-043  
CURRENT APPLICATION NUMBER: US/09/082,279B  
CURRENT FILING DATE: 1998-05-20  
NUMBER OF SEQ ID NOS: 1515  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1488  
LENGTH: 8  
TYPE: PRT  
ORGANISM: HIV-2  
US-09-082-279B-1488

Query Match  
Best Local Similarity 50.0%; Score 26; DB 3; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEIFS 9  
|:|:|  
Db 1 WDVFS 5

RESULT 7  
US-09-315-304B-1642  
Sequence 1642, Application US/09315304B  
Patent No. 6348568  
GENERAL INFORMATION:  
APPLICANT: Barney, S.  
APPLICANT: Guthrie, K.  
APPLICANT: Merutka, G.  
APPLICANT: Anwer, M.  
APPLICANT: Lambert, D.  
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC  
FILE REFERENCE: 7872-052  
CURRENT APPLICATION NUMBER: US/09/315,304B  
PRIOR FILING DATE: 1999-05-20  
PRIOR APPLICATION NUMBER: 09/082,279  
NUMBER OF SEQ ID NOS: 1667  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1642  
LENGTH: 8  
TYPE: PRT  
ORGANISM: HIV-2  
US-09-315-304B-1642

Query Match 50.0%; Score 26; DB 3; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEIFS 9  
|:|:|  
Db 1 WDVFS 5

RESULT 8  
US-09-834-784-1488  
Sequence 1488, Application US/09834784  
Patent No. 6562787  
GENERAL INFORMATION:  
APPLICANT: Barney, Shawn  
APPLICANT: Guthrie, Kelly  
APPLICANT: Merutka, Gene  
APPLICANT: Anwer, Mohamed  
APPLICANT: Lambert, Dennis  
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED  
FILE REFERENCE: 7872-043  
CURRENT APPLICATION NUMBER: US/09/834,784  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: 09/082,279  
NUMBER OF SEQ ID NOS: 1515  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1488  
LENGTH: 8  
TYPE: PRT  
ORGANISM: HIV-2  
US-09-834-784-1488

Query Match 50.0%; Score 26; DB 4; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEIFS 9  
|:|:|  
Db 1 WDVFS 5

RESULT 9  
US-09-350-641C-1643  
Sequence 1643, Application US/09350641C  
Patent No. 6656906  
GENERAL INFORMATION:  
APPLICANT: Barney, S.  
APPLICANT: Guthrie, K.  
APPLICANT: Merutka, G.  
APPLICANT: Anwer, M.  
APPLICANT: Lambert, D.  
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC  
FILE REFERENCE: 7872-067  
CURRENT APPLICATION NUMBER: US/09/350,641C  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: 09/315,304  
PRIOR FILING DATE: 1999-05-20  
PRIOR APPLICATION NUMBER: 09/082,279  
NUMBER OF SEQ ID NOS: 1757  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1643  
LENGTH: 8  
TYPE: PRT  
ORGANISM: HIV-2  
US-09-350-641C-1643

Query Match 50.0%; Score 26; DB 4; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEIFS 9  
|:|:|  
Db 1 WDVFS 5

RESULT 10  
US-08-335-865T-18  
Sequence 18, Application US/08335865T  
Patent No. 6107472  
GENERAL INFORMATION:  
APPLICANT: Stacker, Steven A.; Hovens, Christopher M.,  
APPLICANT: Wilks, Andrew F.  
TITLE OF INVENTION: RECEPTOR-TYPE TYROSINE KINASE-LIKE MOLECULES  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski L.L.P.  
STREET: 666 Fifth Ave  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 720 KB storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: ASCII/Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/335,865T  
FILING DATE: 19-January-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU93/00210  
FILING DATE: 10-May-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PL2358  
FILING DATE: 11-May-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6107472man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD-5277  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3100

TELEFAX: (212) 752-5958  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-357-8657-18

Query Match 48.1%; Score 25; DB 3; Length 7;  
Best Local Similarity 66.7%; Pred. No. 3.8e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLWEL 7  
DB 1 VTLWEL 6

RESULT 11  
US-09-121-211-12  
Sequence 12, Application US/09121211  
Patent No. 675052  
GENERAL INFORMATION:  
APPLICANT: Shinohara, Toshimichi  
APPLICANT: Shingh, Dharendra P.  
APPLICANT: Chylack, Leo T.  
TITLE OF INVENTION: Lens Epithelial Cell Derived Growth  
FILE REFERENCE: B0801/7116  
CURRENT APPLICATION NUMBER: US/09/121,211  
CURRENT FILING DATE: 1998-07-23  
EARLIER APPLICATION NUMBER: U.S. 60/053,549  
EARLIER FILING DATE: 1997-07-23  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 12  
LENGTH: 6  
TYPE: PRP  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)...(6)  
US-09-121-211-12

Query Match 46.2%; Score 24; DB 4; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LWEL 7  
DB 3 LWEL 6

RESULT 12  
US-08-357-642A-4  
Sequence 4, Application US/08357642A  
Patent No. 5837524  
GENERAL INFORMATION:  
APPLICANT: Sina Lev  
APPLICANT: Joseph Schlessinger  
TITLE OF INVENTION: PYK2 RELATED PRODUCTS  
TITLE OF INVENTION: AND METHODS  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/357,642A  
FILING DATE: December 15, 1994  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 209/070  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-357-642A-4

Query Match 46.2%; Score 24; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVLW 5  
DB 4 GVLW 8

RESULT 13  
US-08-357-642A-14  
Sequence 14, Application US/08357642A  
Patent No. 5837524  
GENERAL INFORMATION:  
APPLICANT: Sina Lev  
APPLICANT: Joseph Schlessinger  
TITLE OF INVENTION: PYK2 RELATED PRODUCTS  
TITLE OF INVENTION: AND METHODS  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/357,642A  
FILING DATE: December 15, 1994  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 209/070  
TELECOMMUNICATION INFORMATION:



TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-357-642a-14

Query Match 46.2%; Score 24; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GVLM 5  
Db 4 GVLM 8

RESULT 14  
US-08-460-626-4  
Sequence 4, Application US/08460626  
Patent No. 5837815  
GENERAL INFORMATION:  
APPLICANT: SIMA LEV  
TITLE OF INVENTION: JOSEPH SCHLESSINGER  
TITLE OF INVENTION: PYK-2 RELATED PRODUCTS AND  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,626  
FILING DATE: June 2, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/357,642  
FILING DATE: December 15, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 211/121  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-626-4

Query Match 46.2%; Score 24; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GVLM 5  
Db 4 GVLM 8

RESULT 15  
US-08-460-626-14  
Sequence 14, Application US/08460626  
Patent No. 5837815  
GENERAL INFORMATION:  
APPLICANT: SIMA LEV  
TITLE OF INVENTION: JOSEPH SCHLESSINGER  
TITLE OF INVENTION: PYK-2 RELATED PRODUCTS AND  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,626  
FILING DATE: June 2, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/357,642  
FILING DATE: December 15, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 211/121  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-626-14

Query Match 46.2%; Score 24; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GVLM 5  
Db 4 GVLM 8

Search completed: November 30, 2004, 08:15:58  
Job time : 27 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: November 30, 2004, 08:13:53 : Search time 44 Seconds  
(without alignments)  
80.612 Million cell updates/sec

Title: US-10-008-377a-2

Perfect score: 52

Sequence: 1 GVILWEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 177356

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

- 1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubppa/PCRT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubppa/PCRTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubppa/US09C\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubppa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	14	US-10-008-377-2
2	46	88.5	9	14	US-10-008-377-5
3	32	61.5	10	14	US-09-572-404B-1154
4	31	59.6	9	14	US-10-356-829-22
5	29	55.8	9	15	US-10-149-138-360
6	29	55.8	9	16	US-10-149-138-360
7	29	55.8	10	15	US-10-149-138-92
8	29	55.8	10	15	US-10-149-138-1177
9	29	55.8	10	15	US-10-149-138-2106
10	29	55.8	10	15	US-10-149-138-2707
11	29	55.8	10	16	US-10-149-138-92
12	29	55.8	10	16	US-10-149-138-1177
13	29	55.8	10	16	US-10-149-138-2106

14	29	55.8	10	16	US-10-149-138-2707	Sequence 2707, Ap
15	28	53.8	8	15	US-10-149-138-359	Sequence 359, Ap
16	28	53.8	8	15	US-10-149-138-1176	Sequence 1176, Ap
17	28	53.8	8	15	US-10-149-138-2105	Sequence 2105, Ap
18	28	53.8	8	16	US-10-149-138-359	Sequence 359, Ap
19	28	53.8	8	16	US-10-149-138-1176	Sequence 1176, Ap
20	28	53.8	8	16	US-10-149-138-2105	Sequence 2105, Ap
21	28	53.8	9	9	US-09-884-441-438	Sequence 438, Ap
22	28	53.8	9	10	US-09-907-969-438	Sequence 438, Ap
23	28	53.8	9	10	US-09-827-271-438	Sequence 438, Ap
24	28	53.8	9	14	US-10-198-053-438	Sequence 438, Ap
25	28	53.8	9	14	US-10-245-871-606	Sequence 606, Ap
26	28	53.8	9	15	US-10-149-138-1375	Sequence 1375, Ap
27	28	53.8	9	15	US-10-149-138-3703	Sequence 3703, Ap
28	28	53.8	9	15	US-10-253-286-606	Sequence 606, Ap
29	28	53.8	9	16	US-10-149-138-1375	Sequence 1375, Ap
30	28	53.8	9	16	US-10-149-138-3703	Sequence 3703, Ap
31	28	53.8	10	9	US-09-884-441-421	Sequence 421, Ap
32	28	53.8	10	9	US-09-884-441-433	Sequence 433, Ap
33	28	53.8	10	10	US-09-907-969-421	Sequence 421, Ap
34	28	53.8	10	10	US-09-907-969-433	Sequence 433, Ap
35	28	53.8	10	10	US-09-827-271-421	Sequence 421, Ap
36	28	53.8	10	10	US-09-827-271-433	Sequence 433, Ap
37	28	53.8	10	14	US-10-198-053-431	Sequence 431, Ap
38	28	53.8	10	14	US-10-198-053-433	Sequence 433, Ap
39	28	53.8	10	15	US-10-149-138-1376	Sequence 1376, Ap
40	28	53.8	10	15	US-10-149-138-2000	Sequence 2000, Ap
41	28	53.8	10	16	US-10-149-138-1376	Sequence 1376, Ap
42	28	53.8	10	16	US-10-149-138-2000	Sequence 2000, Ap
43	27	51.9	8	17	US-10-818-770-42	Sequence 42, Ap
44	27	51.9	10	14	US-10-190-082-596	Sequence 596, Ap
45	26	50.0	8	14	US-10-351-641-1643	Sequence 1643, Ap

## ALIGNMENTS

RESULT 1  
US-10-008-377-2  
Sequence 2, Application US/10008377  
Publication No. US2003015701A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
TITLE OF INVENTION: Immunogenic ALK Peptides  
FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008,377  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-2

Query Match  
Best Local Similarity 100.0%; Score 52; DB 14; Length 10;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVILWEIFSL 10  
DB 1 GVILWEIFSL 10

RESULT 2  
US-10-008-377-5  
Sequence 5, Application US/10008377  
Publication No. US2003015701A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
TITLE OF INVENTION: Immunogenic ALK Peptides

FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008,377  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-5

Query Match  
Best Local Similarity 88.5%; Score 46; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLWEIFSL 10  
1 VLWEIFSL 9

RESULT 3  
US-09-572-404B-1154  
Sequence 1154, Application US/09572404B  
Publication No. US20030078374A1  
GENERAL INFORMATION:  
APPLICANT: Proteom Ltd  
TITLE OF INVENTION: Complementary peptide ligands from the human genome  
FILE REFERENCE: Human patent  
CURRENT APPLICATION NUMBER: US/09/572,404B  
CURRENT FILING DATE: 2000-05-17  
NUMBER OF SEQ ID NOS: 4203  
SOFTWARE: Protpatent version 1.0  
SEQ ID NO 1154  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo Sapiens  
FEATURE:  
OTHER INFORMATION: sequence located in Unknown at 860-869 and may interact with sequ  
US-09-572-404B-1154

Query Match  
Best Local Similarity 61.5%; Score 32; DB 10; Length 10;  
Best Local Similarity 83.3%; Pred. No. 77;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEIFSL 10  
1 WEIFSL 6

RESULT 4  
US-10-356-829-22  
Sequence 22, Application US/10356829  
Publication No. US20030155582A1  
GENERAL INFORMATION:  
APPLICANT: BUREAUX, SCOTT RENTON  
APPLICANT: KHANNA, RAJIV  
APPLICANT: SHERIFF, MASTIRA ALISON  
TITLE OF INVENTION: CTL EPITOPES FROM EBV  
FILE REFERENCE: PRC:010USDI  
CURRENT APPLICATION NUMBER: US/10/356,829  
CURRENT FILING DATE: 2003-02-03  
PRIOR APPLICATION NUMBER: 09/462,453  
PRIOR FILING DATE: 2000-01-06  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 22  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Epstein Barr Virus  
US-10-356-829-22

Query Match 59.6%; Score 31; DB 14; Length 9;

Best Local Similarity 83.3%; Pred. No. 1.4e+06;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLWEIF 7  
4 VLWEIF 9

RESULT 5  
US-10-149-138-360  
Sequence 360, Application US/10149138  
Publication No. US20040018971A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Cells, Etseban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
FILE REFERENCE: 2060.0140001  
CURRENT APPLICATION NUMBER: US/10/149,138  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
PRIOR FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 360  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-360

Query Match  
Best Local Similarity 55.8%; Score 29; DB 15; Length 9;  
Best Local Similarity 44.4%; Pred. No. 1.4e+06;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVLWEIFS 9  
1 GVLWEIFMT 9

RESULT 6  
US-10-149-138-360  
Sequence 360, Application US/10149138  
Publication No. US20040121946A9  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Cells, Etseban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
FILE REFERENCE: 2060.0140001  
CURRENT APPLICATION NUMBER: US/10/149,138  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
PRIOR FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 360  
LENGTH: 9

TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-360

Query Match 55.8%; Score 29; DB 16; Length 9;  
Best Local Similarity 44.4%; Pred. No. 1.4e+06;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVLWEIIFS 9  
DB 1 GVTWEIEMT 9

RESULT 7

US-10-149-138-92  
Sequence 92, Application US/10149138  
Publication No. US20040018971A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions  
FILE REFERENCE: 2060.0140001  
CURRENT APPLICATION NUMBER: US/10/149,138  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
PRIOR FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 92  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-92

Query Match 55.8%; Score 29; DB 15; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVLWEIIFS 9  
DB 1 GVTWEIEMT 9

RESULT 8

US-10-149-138-1177  
Sequence 1177, Application US/10149138  
Publication No. US20040018971A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions  
FILE REFERENCE: 2060.0140001  
CURRENT APPLICATION NUMBER: US/10/149,138  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591

PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
PRIOR FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1177  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-1177

Query Match 55.8%; Score 29; DB 15; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVLWEIIFS 9  
DB 1 GVTWEIEMT 9

RESULT 9

US-10-149-138-2106  
Sequence 2106, Application US/10149138  
Publication No. US20040018971A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions  
FILE REFERENCE: 2060.0140001  
CURRENT APPLICATION NUMBER: US/10/149,138  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
PRIOR FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2106  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-2106

Query Match 55.8%; Score 29; DB 15; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVLWEIIFS 9  
DB 1 GVTWEIEMT 9

RESULT 10

US-10-149-138-2707  
Sequence 2707, Application US/10149138  
Publication No. US20040018971A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban

APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
HER2/neu Using Peptide and Nucleic Acid Compositions  
FILE REFERENCE: 2060.0140001  
CURRENT FILING DATE: 2002-06-10  
PCT/US00/33591  
PRIORITY FILING DATE: 2000-12-11  
PRIORITY APPLICATION NUMBER: US 09/458,299  
PRIORITY FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2707  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-2707

Query Match 55.8%; Score 29; DB 16; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVLWEIFS 9  
||:|:|:  
DB 1 GVTWELMT 9

RESULT 11  
US-10-149-138-92  
Sequence 92, Application US/10149138  
Publication No. US20040121946A9  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
HER2/neu Using Peptide and Nucleic Acid Compositions  
FILE REFERENCE: 2060.0140001  
CURRENT FILING DATE: 2002-06-10  
PCT/US00/33591  
PRIORITY FILING DATE: 2000-12-11  
PRIORITY APPLICATION NUMBER: US 09/458,299  
PRIORITY FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 92  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-92

Query Match 55.8%; Score 29; DB 16; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVLWEIFS 9  
||:|:|:  
DB 1 GVTWELMT 9

RESULT 12  
US-10-149-138-1177  
Sequence 1177, Application US/10149138  
Publication No. US20040121946A9

GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
HER2/neu Using Peptide and Nucleic Acid Compositions  
FILE REFERENCE: 2060.0140001  
CURRENT FILING DATE: 2002-06-10  
PCT/US00/33591  
PRIORITY FILING DATE: 2000-12-11  
PRIORITY APPLICATION NUMBER: US 09/458,299  
PRIORITY FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1177  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-1177

Query Match 55.8%; Score 29; DB 16; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVLWEIFS 9  
||:|:|:  
DB 1 GVTWELMT 9

RESULT 13  
US-10-149-138-2106  
Sequence 2106, Application US/10149138  
Publication No. US20040121946A9  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
HER2/neu Using Peptide and Nucleic Acid Compositions  
FILE REFERENCE: 2060.0140001  
CURRENT FILING DATE: 2002-06-10  
PCT/US00/33591  
PRIORITY FILING DATE: 2000-12-11  
PRIORITY APPLICATION NUMBER: US 09/458,299  
PRIORITY FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2106  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-2106

Query Match 55.8%; Score 29; DB 16; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVLWEIFS 9  
||:|:|:  
DB 1 GVTWELMT 9

Db 1 GVTWELMT 9

RESULT 14

US-10-149-138-2707  
 ; Sequence 2707, Application US/10149138  
 ; Publication No. US20040121946A9  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fikes, John  
 ; APPLICANT: Sette, Alessandro  
 ; APPLICANT: Sidney, John  
 ; APPLICANT: Southwood, Scott  
 ; APPLICANT: Chesnut, Robert  
 ; APPLICANT: Celis, Esteban  
 ; APPLICANT: Keogh, Elissa  
 ; TITLE OF INVENTION: Inducing Cellular Immune Responses to  
 ; FILE REFERENCE: 2060.0140001  
 ; CURRENT APPLICATION NUMBER: US/10/149,138  
 ; CURRENT FILING DATE: 2002-06-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/33591  
 ; PRIOR FILING DATE: 2000-12-11  
 ; PRIOR APPLICATION NUMBER: US 09/458,299  
 ; PRIOR FILING DATE: 1999-12-11  
 ; NUMBER OF SEQ ID NOS: 4641  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2707  
 ; LENGTH: 10  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Artificially Synthesized Peptide  
 US-10-149-138-2707

Query Match 55.8%; Score 29; DB 16; Length 10;  
 Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVTWELMT 9  
 Db 1 GVTWELMT 9

RESULT 15

US-10-149-138-359  
 ; Sequence 359, Application US/10149138  
 ; Publication No. US20040018971A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fikes, John  
 ; APPLICANT: Sette, Alessandro  
 ; APPLICANT: Sidney, John  
 ; APPLICANT: Southwood, Scott  
 ; APPLICANT: Chesnut, Robert  
 ; APPLICANT: Celis, Esteban  
 ; APPLICANT: Keogh, Elissa  
 ; TITLE OF INVENTION: Inducing Cellular Immune Responses to  
 ; FILE REFERENCE: 2060.0140001  
 ; CURRENT APPLICATION NUMBER: US/10/149,138  
 ; CURRENT FILING DATE: 2002-06-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/33591  
 ; PRIOR FILING DATE: 2000-12-11  
 ; PRIOR APPLICATION NUMBER: US 09/458,299  
 ; PRIOR FILING DATE: 1999-12-11  
 ; NUMBER OF SEQ ID NOS: 4641  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 359  
 ; LENGTH: 8  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Artificially Synthesized Peptide  
 US-10-149-138-359

Query Match 53.8%; Score 28; DB 15; Length 8;  
 Best Local Similarity 57.1%; Pred. No. 1.4e+06;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVTWELMT 7  
 Db 1 GVTWELMT 7

Search completed: November 30, 2004, 08:31:00  
 Job time : 44 secs

This Page Blank (uspto)



Tue Nov.30 08:49:17 2004

us-10-008-377a-1.closed.rpr

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:41:32; Search time 11.6667 Seconds  
(without alignments)  
82.472 Million cell updates/sec

Title: US-10-008-377A-1

Perfect score: 47

Sequence: 1 SLAMDDLHV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR1:  
2: PIR2:  
3: PIR3:  
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	23	48.9	10 2	E86128
2	19	40.4	10 2	D33098
3	17	36.2	10 2	PH1345
4	16	34.0	6 2	S78764
5	16	34.0	10 2	C26997
6	16	34.0	10 2	A43977
7	15	31.9	7 2	S20446
8	15	31.9	10 2	B46453
9	15	31.9	10 2	A56633
10	15	31.9	10 2	A32543
11	14	29.8	7 2	UN0859
12	14	29.8	7 2	PC0663
13	14	29.8	7 2	A28340
14	14	29.8	10 2	B43590
15	14	29.8	10 2	J00943
16	14	29.8	10 2	A45590
17	14	29.8	10 2	S27873
18	13	27.7	7 2	S55548
19	13	27.7	7 2	A28709
20	13	27.7	7 2	ECMUCR
21	13	27.7	8 2	S66296
22	13	27.7	9 2	B45796
23	13	27.7	9 2	S78426
24	13	27.7	10 2	T40251
25	13	27.7	10 2	T40322
26	13	27.7	10 2	A30823
27	12	25.5	5 2	T10954
28	12	25.5	6 2	A46474
29	12	25.5	7 2	S25266

30	12	25.5	7 2	PT0520
31	12	25.5	7 2	S29735
32	12	25.5	7 2	PC2370
33	12	25.5	8 2	S68802
34	12	25.5	8 2	PT0368
35	12	25.5	8 2	PT0398
36	12	25.5	9 2	E28834
37	12	25.5	9 2	S65913
38	12	25.5	9 2	A33527
39	12	25.5	9 2	S02384
40	12	25.5	9 2	A60108
41	12	25.5	9 2	D48186
42	12	25.5	9 2	S55696
43	12	25.5	9 2	JN0026
44	12	25.5	10 2	A61354
45	12	25.5	10 2	S13224

#### ALIGNMENTS

##### RESULT 1

hypothetical protein Z5903 (imported) - Escherichia coli (strain O157:H7, substrain EDL4)  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: E86128  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamoultis, K.; Apodaca  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A65480; NCID:21074935; PMID:11206551  
A:Accession: E86128  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-10 <STO>  
A:Cross-references: UNIPROT:O8X4E5; GB:AE005174; NID:G12519314; PIDN:AGC9489.1; GSPDB:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z5903

Query Match 48.9%; Score 23; DB 2; Length 10;  
Best Local Similarity 57.1%; Pred. No. 1.4e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 MLDDLHV 10  
DB 1 MIDILQV 7

RESULT 2  
D33098  
214K exantigen (version 1) - malaria parasite (Plasmodium falciparum) (fragments)  
C:Species: Plasmodium falciparum  
C:Date: 24-Aug-1990 #sequence\_revision 24-Aug-1990 #text\_change 09-Jun-2000  
C:Accession: D33098  
R:Nichols, J.H.; Hager, L.P.  
submitted to the Protein Sequence Database, May 1990  
A:Reference number: A33098  
A:Accession: D33098  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <NIC>

Query Match 40.4%; Score 19; DB 2; Length 10;  
Best Local Similarity 80.0%; Pred. No. 9.1e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAMLD 6  
DB 2 LAMLD 6

## RESULT 3

PH1345

Ig heavy chain DJ region (clone C100-94) - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999

C/Accession: PH1345

A/Species: Man; R.; Gallil, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992

A/Title: Predominance of fetal type DJH joining in young children with B precursor lymph

A/Reference number: PH1302; MUID:93094761; PMID:1460419

A/Accession: PH1345

A/Molecule type: DNA

A/Residues: 1-10 &lt;MAS&gt;

A/Note: the authors translated the stop codon for residue 4 as X

C/Keywords: heterotrimer; immunoglobulin

QY

3 AMDL 8

DB

1 AMVXLL 6

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

DB

1 MDL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

DB

1 MDL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

DB

1 MDL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

DB

1 MDL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

DB

1 MDL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

DB

1 MDL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

A/Note: sequence extracted from NCBI backbone (NCBIP:60243)  
F:1-10/Domain: signal sequence (fragment) #status predicted <SIG>

Query Match 31.9%; Score 15; DB 2; Length 10;  
Best Local Similarity 33.3%; Pred. No. 6e+03;  
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 DDLHV 10  
|:|:  
1 MQLFHL 6

RESULT 9  
A56633  
neomysuppressin - flesh fly (Sarcophaga bullata)

N/Alternate names: Neb-MS  
C/Species: Sarcophaga bullata  
C/Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
C/Accession: A56633  
R/Fonagy, A.; Schoofs, L.; Proost, P.; Van Damme, J.; Bueds, H.; De Loof, A.  
Comp. Biochem. Physiol. C 102, 239-245, 1992  
A/Title: Isolation, primary structure and synthesis of neomysuppressin, a myoinhibiting  
A/Reference number: A56633; PMID:93047886; PMID:1356537  
A/Accession: A56633  
A/Molecule type: protein  
A/Residues: 1-10 <FON>  
A/Cross-references: UNIPROT:P61850  
A/Experimental source: head  
A/Note: sequence extracted from NCBI backbone (NCBIP:119072)  
C/Keywords: amidated carboxyl end; neuropeptide  
F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 31.9%; Score 15; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 6e+03;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 DDLHV 10  
|:|:  
2 DVDHV 6

RESULT 10  
A32543  
cardioexcitatory neuropeptide - desert locust  
C/Species: Schistocerca gregaria (desert locust)  
C/Date: 20-Dec-1989 #sequence\_revision 20-Dec-1989 #text\_change 09-Jul-2004  
C/Accession: A32543  
R/Robb, S.; Packman, L.C.; Evans, P.D.  
Biochem. Biophys. Res. Commun. 160, 850-856, 1989  
A/Title: Isolation, primary structure and bioactivity of Schistocerca gregaria, a FMRF-amide  
A/Reference number: A32543; PMID:89246543; PMID:2719702  
A/Accession: A32543  
A/Molecule type: Preliminary  
A/Status: Preliminary  
A/Residues: 1-10 <ROB>  
A/Cross-references: UNIPROT:P38553  
C/Suprafamily: unassigned animal peptides  
C/Keywords: amidated carboxyl end; neuropeptide  
F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 31.9%; Score 15; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 6e+03;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 DDLHV 10  
|:|:  
2 DVDHV 6

RESULT 11  
JN0859  
peptidyl-dipeptidase A inhibitory peptide C105 - striped bonito  
C/Species: Sarda orientalis (striped bonito)

C/Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 07-May-1999  
C/Accession: JN0859  
R/Matsumura, N.; Fujii, M.; Takeda, Y.; Shimizu, T.  
Biochim. Biophys. Acta 1174, 174-174, 1993  
A/Title: Isolation and characterization of angiotensin I-converting enzyme inhibitory P  
A/Reference number: JN0859; PMID:9480036; PMID:7764272  
A/Accession: JN0859  
A/Molecule type: protein  
A/Residues: 1-7 <MAT>  
A/Experimental source: intestine  
A/Comment: The carboxyl-terminus is essential for the protein's expression of angiotensin  
C/Suprafamily: bradykinin-potentiating peptide  
C/Keywords: angiotensin-converting enzyme inhibitor

Query Match 29.8%; Score 14; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMD 6  
|:|:  
1 SVAKLE 6

RESULT 12

P00663  
membrane protein - porcine epidemic diarrhea virus (isolate Belgian CV777) (fragment)  
C/Species: porcine epidemic diarrhea virus  
C/Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 08-Oct-1999  
C/Accession: P00663  
R/Bridgen, A.; Duarte, M.; Tobler, K.; Laude, H.; Ackermann, M.  
J. Gen. Virol. 74, 1795-1804, 1993  
A/Title: Sequence determination of the nucleocapsid protein gene of the porcine epidemic  
A/Reference number: J02191; PMID:93389433; PMID:8397280  
A/Accession: P00663  
A/Molecule type: mRNA  
A/Residues: 1-7 <BR>  
A/Cross-references: GB:Z14976; NID:G311650; PIDN:CAAT8699.1; PID:9584083  
C/Comment: This virus is coronavirus related to human coronavirus 229E.  
C/Keywords: membrane protein

Query Match 29.8%; Score 14; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 LHV 10  
|:|:  
3 VLHV 6

RESULT 13

A28340  
myomodulin - California sea hare  
C/Species: Aplysia californica (California sea hare)  
C/Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
C/Accession: A28340  
R/Cropper, E.C.; Tenenbaum, R.; Kols, M.A.G.; Kumpfmann, I.; Weiss, K.R.  
Proc. Natl. Acad. Sci. U.S.A. 84, 5483-5486, 1987  
A/Title: Myomodulin: A bioactive neuropeptide present in an identified cholinergic buccal  
A/Reference number: A28340; PMID:87261010; PMID:3474664  
A/Accession: A28340  
A/Molecule type: protein  
A/Residues: 1-7 <CRO>  
A/Cross-references: UNIPROT:P15513

Query Match 29.8%; Score 14; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMDL 7  
|:|:  
2 MSMRL 7

## RESULT 14

B43590

pili type A66 - Aeromonas hydrophila (fragment)

C:Species: Aeromonas hydrophila

C:Date: 12-Jan-1993 #sequence\_revision 12-Jan-1993 #text\_change 17-Nov-2000

C:Accession: B43590

R:Hokama, A.; Iwanaga, M.

Infect. Immun. 59, 3478-3483, 1991

A:Title: Purification and characterization of Aeromonas sobria pili, a possible colonizer

A:Reference number: A43590; PMID:91372953; PMID:1680105

A:Accession: B43590

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 &lt;HOK&gt;

C:Superfamily: Vibrio cholerae prepilin-like 17.3k protein type 4

Query Match

Best Local Similarity 29.8%; Score 14; DB 2; Length 10;

Matches 2; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAMDLHLV 10

DB 1 MTLIELVIV 9

## RESULT 15

J00943

hypothetical 1.3k protein - infectious bursal disease virus (strain 52/70)

C:Species: infectious bursal disease virus

C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004

C:Accession: J00943

R:Baylis, C.D.; Spies, U.; Shaw, K.; Peters, R.W.; Papageorgiou, A.; Mueller, H.; Bourc

U. Gen. Virol. 71, 1303-1312, 1990

A:Title: A comparison of the sequences of segment A of four infectious bursal disease vi

A:Reference number: J00941; PMID:90278420; PMID:2161902

A:Accession: J00943

A:Status: translation not shown

A:Molecule type: genomic RNA

A:Residues: 1-10 &lt;BAY&gt;

A:Cross-references: UNIPROT:Q82625; GB:000869; NID:9221036; PIDN:BAA00743.1; PID:dl00120

Query Match

Best Local Similarity 29.8%; Score 14; DB 2; Length 10;

Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LAMDLHLV 9

DB 3 LLLLRVH 10

Search completed: November 30, 2004, 08:14:25  
 Job time : 12.6667 secs

```

32 15 31.9 10 2 Q7S225 07s225 neuropeptidase
33 15 31.9 10 2 Q6H101 06h101 canis fam1
34 15 31.9 10 2 P74843 p74843 streptomyc
35 15 31.9 10 2 P82588 p82588 streptococ
36 15 31.9 10 2 Q9YXC3 q9yxc3 hepatitis b
37 14 29.8 7 2 Q9YQ10 q9yq10 transmissib
38 14 29.8 8 2 Q7Z6D0 q7z6d0 homo sapien
39 14 29.8 8 2 Q6Z246 q6z246 bacillus ps
40 14 29.8 9 1 P811_LYCES p811_lyces
41 14 29.8 9 2 Q7R8X5 q7r8x5 lycopterisico
42 14 29.8 9 2 Q8WS58 q8ws58 anthocidari
43 14 29.8 9 2 P83539 p83539 lactobacill
44 14 29.8 9 2 Q91LK6 q91lk6 macaca neme
45 14 29.8 9 2 Q61754 q61754 cynops pyrtr

ALIGNMENTS

RESULT 1
Q8X4E5 PRELIMINARY: PRT; 10 AA.
ID Q8X4E5
AC Q8X4E5
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Z5903 protein.
GN OrderedLocustNames=z5903;
OS Escherichia coli O157:H7;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL93 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206553;
RA Pena N.T., Plunkett G., Iff, Bertrand V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Pofel G., Hackett U., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouzis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blither F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RT Nature 409:529-533(2001).
DR EMBL; AEO05662; AAG59489.1; -.
DR PIR; E86128; B86128.
KW Complete proteome.
SQ SEQUENCE 10 AA; 1161 MW; 98AC3BD2D6D7205A CRC64;

Query Match 48.9%; Score 23; DB 2; Length 10;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0.

QY 4 MEDLIHV 10
DB 1 MIDLIQV 7

RESULT 2
Q37854 PRELIMINARY: PRT; 8 AA.
ID Q37854
AC Q37854;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE Coliphage gene of unknown function, 5'end. (Fragment).
OS Bacteriophage R17.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;
OC Levivirus
OX NCBI_TaxID=12026;
[1]
RN
RP SEQUENCE FROM N.A.
RC MEDLINE=73224967; PubMed=4352721;
RX

```

RA Rensing U.F.E.;  
 RT "A sequence of seventy-three nucleotides from the coliphage R17  
 genome.";  
 RL Biochem. J. 131:593-604(1973).  
 DR EMBL: M24820; AAA72755.1; -.  
 FT NON TER 8  
 SQ SEQUENCE 8 AA; 969 MW; ECB45412C1E72726 CRC64;

Query Match 42.6%; Score 20; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 LHMV 10  
 Db 2 LHMV 5

RESULT 3  
 O51594 PRELIMINARY; PRT; 8 AA.  
 AC O51594;  
 DT 01-NOV-1996 (TREMURel. 01, Created)  
 DT 01-NOV-1996 (TREMURel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMURel. 24, Last annotation update)  
 DE CoxB protein (Fragment).  
 OS Plasmid ColV2-K94.  
 OG Plasmid.  
 OC Plasmids.  
 OX NCBI\_TaxID=2458;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=6623772; PubMed=2423502;  
 RA Weber P.C. Palchaudhuri S.;  
 RT "Incompatibility repressor in a RepA-like replicon of the IncFI  
 plasmid ColV2-K94.";  
 RL J. Bacteriol. 166:1106-1121(1986).  
 DR EMBL: M13472; AAA23194.1; -.  
 KW Plasmid.  
 FT NON TER 1  
 SQ SEQUENCE 8 AA; 998 MW; 0F37205AA7341ED7 CRC64;

Query Match 38.3%; Score 18; DB 2; Length 8;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 2 LAMDL 8  
 Db 1 LORIDL 7

RESULT 4  
 RE42\_LITRU STANDARD; PRT; 9 AA.  
 ID RE42\_LITRU  
 AC P82075; P82093;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Rubellidin 4.2/4.3.  
 OS Litorea rubella (Desert tree frog).  
 CC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;  
 CC Pelodyadinae; Litorea.  
 OX NCBI\_TaxID=104895;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC Tissue-Skin secretion;  
 RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,  
 RA Tyler M.J., Wallace J.C.;  
 RT "The structure of new peptides from the Australian red tree frog  
 'Litorea rubella'. The skin peptide profile as a probe for the study  
 of evolutionary trends of amphibians.";  
 RL Aust. J. Chem. 49:955-963(1996).  
 RN [2]

RP SEQUENCE.  
 RC Tissue-Skin secretion;  
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;  
 RT "Peptides from the skin glands of the Australian tree frog  
 'Litorea rubella'. Comparison with the skin peptides from Litorea  
 rubella.";  
 RL Aust. J. Chem. 52:639-645(1999).  
 CC -1- FUNCTION: Shows neither neuropeptide activity nor antibiotic  
 activity.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.  
 CC -1- PTM: Rubellidin 4.2 seems to differ from Rubellidin 4.3 by its C-  
 terminal amidation.  
 CC -1- MASS SPECTROMETRY: MW=883; METHOD=FAE; RANGE=1-9; NOTE=Ref. 1.  
 KW Amidation; Amphibian defense peptide; Direct protein sequencing.  
 FT MOD RES 9  
 SQ SEQUENCE 9 AA; 984 MW; 2C2D7205AA72728 CRC64;

Query Match 38.3%; Score 18; DB 1; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 4 MLDL 8  
 Db 3 LIDL 7

RESULT 5  
 Q7RD8 PRELIMINARY; PRT; 9 AA.  
 ID Q7RD8  
 AC Q7RD8;  
 DT 01-MAR-2004 (TREMURel. 26, Created)  
 DT 01-MAR-2004 (TREMURel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMURel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN Name=Pt0516;  
 OS Plasmodium yoelii yoelii.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=73239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=17XNL;  
 RX PubMed=1236865;  
 RA Carlton J.M., Anguoli S.V., Suh B.B., Kocil T.W., Pettea M.,  
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koc H.L.,  
 RA Peterson J.D., Pop M., Kosack D.S., Schumway M.F., Bidwell S.L.,  
 RA Shalom S.J., van Aken S.E., Riedmiller S.B., Feldblum T.V.,  
 RA Cho J.K., Quackenbush J., Sedegah M., Shoaihi A., Cummings L.M.,  
 RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,  
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Valdivia A.B.,  
 RA van Lin H., Janse C.J., Waters A.P., Smith H.O., White O.R.,  
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,  
 RA Carucci D.J.;  
 RT "Genome sequence and comparative analysis of the model rodent malaria  
 parasite Plasmodium yoelii yoelii.";  
 RL Nature 419:512-519(2002).  
 CC -1- CAUTION: The sequence shown here is derived from an  
 EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is  
 preliminary data.  
 DR EMBL: AAB01001760; EAA17550.1; -.  
 KW Hypothetical protein  
 SQ SEQUENCE 9 AA; 1116 MW; 500A3B1337272447 CRC64;

Query Match 38.3%; Score 18; DB 2; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 4 MLDL 8  
 Db 1 MLDL 5

RESULT 6

Q66205  
ID 066205 PRELIMINARY; PRT; 7 AA.  
AC 066205;  
DT 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE Membrane protein (1 is 3rd base in codon) (Fragment).  
OC Transmissible gastroenteritis virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
OC Coronaviridae; Coronavirus; Group 1 species.  
NCBI\_TaxID=11149;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ES772/70;  
RA MEDLINE=88216185; PubMed=2835592;  
RA Britton P., Carmenes R.S., Page K.W., Garves D.J., Parra F.;  
RT "Sequence of the nucleoprotein gene from a virulent British field  
RT isolate of transmissible gastroenteritis virus and its expression in  
RT Saccharomyces cerevisiae."  
RL Mol. Microbiol. 2:89-99(1988).  
DR EMBL: Y00542; CAA68606.1; -.  
FT NON TER 1  
SQ SEQUENCE 7 AA; 869 MW; 72C691E727233B70 CRC64;

Query Match 36.2%; Score 17; DB 2; Length 7;  
Best Local Similarity 75.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 7 LHM 10  
DB 3 LHM 6

RESULT 7  
Q9TT78 PRELIMINARY; PRT; 8 AA.  
ID 09TT78  
AC 09TT78;  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE Thymidylate synthase (Fragment).  
OS Caris familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21015404; PubMed=11130975;  
RA Brouillette J.A., Andrew J.R., Venta P.J.;  
RL Mamm. Genome 11:1079-1086(2000).  
DR EMBL: AF202073; AAF20918.1; -.  
FT NON TER 1  
SQ SEQUENCE 8 AA; 899 MW; 6731A1B059CA867 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 8;  
Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 LHM 9  
DB 3 DFLH 6

RESULT 8  
O70SM2 PRELIMINARY; PRT; 9 AA.  
ID 070SM2  
AC 070SM2;  
DT 05-JUL-2004 (TREMblrel. 27, Created)  
DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)

DE Hypermethylated in cancer 1 (Fragment).  
GN Name=HIC1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Pinte S., Guerardel C., Deltour S., Godwin A.R., Lepince D.;  
RT "Identification of a second G-C-rich promoter conserved in the human,  
RT murine and rat tumor suppressor genes HIC1."  
RL Oncogene 23:4023-4031(2004).  
DR EMBL: AF550616; CAD79467.1; -.  
FT NON TER 9  
SQ SEQUENCE 9 AA; 964 MW; 5B5E6DB1681AA7 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 MDDL 8  
DB 1 MDTM 5

RESULT 9  
Q8GIZ6 PRELIMINARY; PRT; 9 AA.  
ID 08GIZ6  
AC 08GIZ6;  
DT 01-MAR-2003 (TREMblrel. 23, Created)  
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)  
DE Glyceralddehyde 3-phosphate dehydrogenase (EC 1.2.1.12)  
DE (Fragment).  
GN Name=gap;  
OS Lactobacillus delbrueckii (subsp. lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
OC Lactobacillus.  
OC NCBI\_TaxID=29397;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCC89;  
RA Bourbiquet A.A., Mollet B.;  
RT "Purification and characterization of the 3-phosphoglycerate kinase  
RT from the thermophile Lactobacillus delbrueckii subsp. lactis."  
RL Int. Dairy J. 12:723-728(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCC89;  
RA Bourbiquet A.A.;  
RT "Molecular insights into the metabolism and physiology of the lactic  
RT acid bacterium Lactobacillus delbrueckii subsp. lactis."  
RL Thesis (2000), Department of Molecular Microbiology, Biozentrum der  
RL Universitaet Basel, (PhD work conducted at the Nestle Research Center,  
RL Lausanne), Switzerland.  
DR EMBL: AJ315554; CAD56494.1; -.  
DR GO: GO:0004365; F:Glyceralddehyde-3-phosphate dehydrogenase (p. .; IEA.  
DR GO: GO:0016491; F:oxidoreductase activity; IEA.  
FT NON TER 1  
SQ SEQUENCE 9 AA; 1071 MW; 94ABAD9C1E72721 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LHM 9  
DB 3 LHM 5

RESULT 10

OSQZAT7 PRELIMINARY; PRT; 9 AA.

AC OSQZAT7.  
 DT 01-MAY-2000 (TREMUREL. 13, Created)  
 DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMUREL. 13, Last annotation update)  
 DE TARC protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBI\_TaxID=10090;  
 RP SEQUENCE FROM N.A.  
 RA Gorski K., Huang X., Tseng S.-Y., Rattis F., Pardoll D., Tsuchiya H.;  
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF192527; AF04844.1; -  
 FT NON TER 9  
 SQ SEQUENCE 9 AA; 1104 MW; 621E272686D735A4 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SLAML 5  
 DB 3 SLOWL 7

RESULT 11

OSUTD7 PRELIMINARY; PRT; 9 AA.

AC OSUTD7.  
 DT 01-MAR-2002 (TREMUREL. 20, Created)  
 DT 01-MAR-2002 (TREMUREL. 20, Last sequence update)  
 DT 01-OCT-2002 (TREMUREL. 22, Last annotation update)  
 DE Vpu protein.  
 GN Name=vpu;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroviruses; Retroviridae; Lentivirus.  
 RN NCBI\_TaxID=11676;  
 RP SEQUENCE FROM N.A.  
 RA Novitsky V., Smith U.R., Gilbert P., McLane M.F., Chigwedere P.,  
 RA Williamson C., Ndung'u T., Klein I., Chang S.-Y., Peter T., Thior I.,  
 RA Foley B.T., Goletskwe S., Rybak N., Gaseitlswe S., Vanberg F.,  
 RA Marink R., Lee T.-H., Essex M.;  
 RT "Human immunodeficiency virus type 1 subtype C molecular phylogeny:  
 RT consensus sequence for an AIDS vaccine design";  
 RL J. Virol. 76:5435-5451(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Novitsky V.A., McLane M.F., Chigwedere P., Ndung'u T., Klein I.,  
 RA Chang S.-Y., Peter T., Thior I., Rybak N., Gaseitlswe S., Vanberg F.,  
 RA Marink R., Lee T.-H., Essex M.;  
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF443091; AL34712.1; -  
 SQ SEQUENCE 9 AA; 1102 MW; 188BD40B17272440 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 MDL 8  
 DB 1 MINL 5

RESULT 12

CAD79467 PRELIMINARY; PRT; 9 AA.

ID CAD79467.  
 AC CAD79467.  
 DT 01-JUN-2004 (TREMUREL. 27, Created)

DT 01-JUN-2004 (TREMUREL. 27, Last sequence update)  
 DT 01-JUN-2004 (TREMUREL. 27, Last annotation update)  
 DE Hypermethylated in cancer 1 (Fragment).  
 GN HIC1.  
 OS Homo sapiens (human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 RN NCBI\_TaxID=9606;  
 RP SEQUENCE FROM N.A.  
 RA Tissue=Mammary gland;  
 RX PubMed=15007385;  
 RA Pinte S., Guerardel C., Delcoul S., Godwin A.K., Lepine D.;  
 RT "Identification of a second G-C-rich promoter conserved in the human,  
 RT murine and rat tumor suppressor genes HIC1.";  
 RL Oncogene 23:4023-4031(2004).  
 DR EMBL; AJ550616; CAD79467.1; -  
 FT NON TER 9  
 SQ SEQUENCE 9 AA; 964 MW; 5B5E6DD81681AA7 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 MDL 8  
 DB 1 MDLTM 5

RESULT 13

FARP MANSE STANDARD; PRT; 10 AA.

AC P18523.  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE FMRamide-like neuropeptide.  
 OS Manduca sexta (tobacco hawkmoth) (Tobacco hornworm).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;  
 CC Sphingidae; Sphinginae; Manduca.  
 RN NCBI\_TaxID=7130;  
 RP SEQUENCE.  
 RA MEDLINE=91045350; PubMed=2235684;  
 RA Kingan T.G., Teplov D.B., Phillips J.M., Riehm J.P., Rao K.R.,  
 RA Hildebrand J.G., Homberg U., Kammer A.E., Jardine I., Griffin P.R.,  
 RA Hunt D.F.;  
 RT "A new peptide in the FMRamide family isolated from the CNS of the  
 RT hawkmoth, Manduca sexta.";  
 RL Peptides 11:849-856(1990).  
 CC -1- FUNCTION: Increases the force of neurally evoked contractions in  
 CC the major power-producing flight muscles, the dorsal longitudinal  
 CC muscles and so is likely to play a role in sustaining or promoting  
 CC flight behavior patterns.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)  
 CC family.  
 DR PIR: A43977; A43977.  
 RX Amidation: Direct protein sequencing; Neuropeptide;  
 KM Pyrrolidone carboxylic acid. Pyrrolidone carboxylic acid.  
 FT MOD\_RES 1  
 FT MOD\_RES 10  
 FT MOD\_RES 10  
 SQ SEQUENCE 10 AA; 1247 MW; D3C4523D5B1F2D2 CRC64;

Query Match 34.0%; Score 16; DB 1; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+04;  
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 6 DLH 9  
 DB 2 DVH 5



## RESULT 14

Q9TS43 PRELIMINARY; PRT; 10 AA.  
 AC Q9TS43;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE OESTRADIOL-RECEPTOR-PI peptide (fragment).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9129128; PubMed=2064608;  
 RA Thole H.H., Jungblut P.W., Jakob F.;  
 RT "The proton-driven dissociation of oestradiol-receptor dimers as a  
 RT preparative tool. Isolation of a 32 kDa fragment from porcine uteri  
 RT and assignment of C-terminal origin by partial sequencing.";  
 RL Biochem. J. 276:709-714(1991).  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 10 AA; 1240 MW; D31AD274405691F2 CRC64;

## Query Match

Best Local Similarity 34.0%; Score 16; DB 2; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 2.8e+04;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 DLHIV 10  
 :|||:  
 Db 1 ELYHM 5

## RESULT 15

039949 PRELIMINARY; PRT; 10 AA.  
 ID 039949  
 AC 039949; 039950; 039953; 039954; 039955;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE El protein (fragment).  
 OS GB virus C/Hepatitis G virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC GSV-C/HGV group.  
 OX NCBI\_TaxID=54290;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97368412; PubMed=9225026;  
 RA Smith D.B., Cuccannu N., Davidson P., Jarvis L.M., Mokili J.L.,  
 RA Hamid S., Ludlam C.A., Simmonds P.;  
 RT "Discrimination of hepatitis G virus/GBV-C geographical variants by  
 RT analysis of the 5' non-coding region."  
 RL J. Gen. Virol. 78:1533-1542(1997).  
 DR EMBL; AF003167; AAC57978.1; -  
 FT NON\_TER 10  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1083 MW; CC88FE27272732 CRC64;

## Query Match

Best Local Similarity 34.0%; Score 16; DB 2; Length 10;  
 Best Local Similarity 55.6%; Pred. No. 2.8e+04;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAMDLHIV 10  
 :|||:  
 Db 1 MAVLTLIV 9

Search completed: November 30, 2004, 08:13:39  
 Job time : 80.3333 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:34:52; Search time 55.3333 Seconds  
(without alignments)  
64.831 Million cell updates/sec

Title: US-10-008-377A-1  
Perfect score: 47  
Sequence: 1 SLAMDLHLV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 444336

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: A\_Genesep23Sep04:\*

- 1: genesep1980s:\*
- 2: genesep1990s:\*
- 3: genesep2000s:\*
- 4: genesep2001s:\*
- 5: genesep2002s:\*
- 6: genesep2003as:\*
- 7: genesep2003bs:\*
- 8: genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	10	6	AAO22983 Human p28
2	43	91.5	9	6	AAO22985 Human p28
3	39	83.0	9	6	AAO22986 Human p28
4	26	55.3	9	6	ABU03234 Human exp
5	26	55.3	10	7	AAW45584 Peptide f
6	26	55.3	10	7	ADK24124 Human 98P
7	26	55.3	10	7	ADK24937 Human 98P
8	25	53.2	9	7	ADK65168 Human VEG
9	25	53.2	10	4	AAE65116 PSA immun
10	25	53.2	10	4	AAE68173 Saccharom
11	25	53.2	10	4	AAE68172 Saccharom
12	24	51.1	9	6	AAE22055 Human can
13	24	51.1	9	6	ABU03262 Human exp
14	24	51.1	10	5	AAU94381 Human nov
15	24	51.1	10	7	ADK24903 Human 98P
16	24	51.1	10	8	ADJ83977 Human asp
17	23	48.9	9	5	AAE31277 Human mag
18	23	48.9	10	2	AAV40366 Human mag
19	23	48.9	10	2	AAV5317 HIV-1 net
20	23	48.9	10	2	AAV5317 HIV-1 net
21	23	48.9	10	2	AAV5317 HIV-1 net
22	23	48.9	10	2	AAV5317 HIV-1 net
23	22.5	47.9	10	5	ABU17292 Zinc tran
24	22.5	47.9	10	5	ABU17024 Zinc tran
25	22	46.8	7	2	AAW46002 Peptide #

26	22	46.8	7	4	AAW98287 Human pep
27	22	46.8	7	8	ADH68228 Human G-P
28	22	46.8	9	2	AAE73813 Antigen f
29	22	46.8	9	2	AAE47722 Immunogen
30	22	46.8	9	2	AAE55465 HLA bindi
31	22	46.8	9	4	AAE75668 HLA class
32	22	46.8	9	5	AAE6083 Human glu
33	22	46.8	9	5	AAE28729 Human GAS
34	22	46.8	9	5	AAE9426 Human nov
35	22	46.8	9	5	AAE9426 Human nov
36	22	46.8	9	5	AAE9426 Human nov
37	22	46.8	9	5	AAE9426 Human nov
38	22	46.8	9	5	AAE9426 Human nov
39	22	46.8	9	5	AAE9426 Human nov
40	22	46.8	9	5	AAE9426 Human nov
41	22	46.8	9	5	AAE9426 Human nov
42	22	46.8	9	5	AAE9426 Human nov
43	22	46.8	9	5	AAE9426 Human nov
44	22	46.8	9	5	AAE9426 Human nov
45	22	46.8	9	5	AAE9426 Human nov

## ALIGNMENTS

RESULT 1  
AAO22983 standard; peptide; 10 AA.

AAO22983;  
17-SEP-2003 (first entry)

Human p28-89 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.

HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL; cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human; oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy; t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma; p280-89.

OS Homo sapiens.  
XX WO2003042243-A2.  
XX 22-MAY-2003.  
XX 14-NOV-2002; 2002WC-EP012764.  
XX 15-NOV-2001; 2001US-00008377.  
XX (MAST-) IST NAZ STUDIO E CURA DEI TUMORI.  
XX (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX Gambacorti-Passerini C, Passoni L;  
XX WPI; 2003-441791/41.  
XX New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
XX useful for preparing a composition for treating ALK-positive lymphoma,  
XX neuroblastoma or ALK-expressing neoplasia.  
XX Claim 1; Page 3; 33p; English.

The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell lines. More than 50% of ALCL cases possess a t(2;5) chromosomal translocation that leads to the expression of the NPM/ALK fusion protein which forms a potent oncogene when constitutively activated. Translocated ALK is a widely expressed tumour-associated antigen characteristic of ALK

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p280-89  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 CC  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 47; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 SLAMF1LHV 10  
 DB 1 SLAMF1LHV 10

RESULT 2  
 AAO22985

ID AAO22985 standard; peptide; 9 AA.

AC AAO22985;

DT 17-SEP-2003 (first entry)

DE Human p281-89 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.

XX HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
 XX cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;  
 XX oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
 XX t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
 XX p281-89.

XX Homo sapiens.

XX WO2003042243-A2.

XX 22-MAY-2003.

XX 14-NOV-2002; 2002MO-EP012764.

XX 15-NOV-2001; 2001US-00008377.

XX (NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.

XX (SUUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Gambacorti-Passerini C, Passoni L;

XX WPI; 2003-441791/41.

PT New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
 PT useful for preparing a composition for treating ALK-positive lymphoma,  
 PT neuroblastoma or ALK-expressing neoplasia.

XX Claim 1; Page 7; 33pp; English.

XX The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma  
 CC kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a  
 CC cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase  
 CC which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase  
 CC fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell  
 CC lines. More than 50% of ALCL cases possess a t(2;5) chromosomal  
 CC translocation that leads to the expression of the NPM/ALK fusion protein  
 CC which forms a potent oncogene when constitutively activated. Translocated  
 CC ALK is a widely expressed tumour-associated antigen characteristic of ALK  
 CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion

CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p281-89  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 CC  
 SQ Sequence 9 AA;

Query Match 91.5%; Score 43; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2 SLAMF1LHV 10  
 DB 1 SLAMF1LHV 9

RESULT 3  
 AAO22986

ID AAO22986 standard; peptide; 9 AA.

AC AAO22986;

DT 17-SEP-2003 (first entry)

DE Human p282-90 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.

XX HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
 XX cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;  
 XX oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
 XX t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
 XX p282-90.

XX Homo sapiens.

XX WO2003042243-A2.

XX 22-MAY-2003.

XX 14-NOV-2002; 2002MO-EP012764.

XX 15-NOV-2001; 2001US-00008377.

XX (NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.

XX (SUUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Gambacorti-Passerini C, Passoni L;

XX WPI; 2003-441791/41.

PT New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
 PT useful for preparing a composition for treating ALK-positive lymphoma,  
 PT neuroblastoma or ALK-expressing neoplasia.

XX Claim 1; Page 7; 33pp; English.

XX The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma  
 CC kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a  
 CC cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase  
 CC which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase  
 CC fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell  
 CC lines. More than 50% of ALCL cases possess a t(2;5) chromosomal  
 CC translocation that leads to the expression of the NPM/ALK fusion protein  
 CC which forms a potent oncogene when constitutively activated. Translocated  
 CC ALK is a widely expressed tumour-associated antigen characteristic of ALK  
 CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p282-90

CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
XX Sequence 9 AA;  
SQ

Query Match 83.0%; Score 39; DB 6; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMLDLHLV 10  
| | | | |  
| | | | |  
Db 1 AMLDLHLV 8

RESULT 4  
ABU03234  
ID ABU03234 standard; protein; 9 AA.  
XX  
XX ABU03234;  
XX  
XX  
XX 29-JAN-2003 (first entry)  
XX  
XX Human expressed protein tag (EPT) #14.  
XX  
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX protease; protease inhibitor; transporter; cytoskeletal protein;  
XX receptor; transcription factor; cancer; MHC;  
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
XX Homo sapiens.  
XX  
XX MO200278524-A2.  
XX  
XX 10-OCT-2002.  
XX  
XX 28-MAR-2002; 2002WO-US009671.  
XX  
XX 28-MAR-2001; 2001US-0279495P.  
XX 21-MAY-2001; 2001US-0292544P.  
XX 08-AUG-2001; 2001US-0310801P.  
XX 01-OCT-2001; 2001US-0326370P.  
XX 04-DEC-2001; 2001US-0336780P.  
XX 20-FEB-2002; 2002US-0358985P.  
XX  
XX (ZYCO-) ZYCO INC.  
XX  
XX Chicx RM, Tomlinson AJ, Urban RG;  
XX  
XX WPI; 2003-040607/03.  
XX  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.  
XX  
XX Claim 10; SEQ ID NO 14; 134pp; English.

CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 9 AA;  
SQ

Query Match 55.3%; Score 26; DB 6; Length 9;  
Best Local Similarity 71.4%; Pred. No. 1.7e+06;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AMLDLHLV 9  
| | | | |  
| | | | |  
Db 1 ALDLHLV 7

RESULT 5  
AAW45584  
ID AAW45584 standard; peptide; 10 AA.  
XX  
XX AAW45584;  
XX  
XX  
XX 03-JUN-1998 (first entry)  
XX  
XX Peptide fragment of leptin (ob 106-115) that modulates body weight.  
XX  
XX Leptin; obesity; body weight; diabetes; energy; metabolic disorder;  
XX ob protein.  
XX  
XX Homo sapiens.  
XX  
XX MO9746585-A2.  
XX  
XX 11-DEC-1997.  
XX  
XX 04-JUN-1997; 97WO-EP02968.  
XX  
XX 06-JUN-1996; 96GB-00011775.  
XX 05-SEP-1996; 96GB-00018540.  
XX 20-FEB-1997; 97GB-00003493.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM PLC.  
XX  
XX Albaranzanji KA, Arch JR, Camilleri P, Neville WA;  
XX  
XX WPI; 1998-042120/04.  
XX  
XX Peptide fragments of leptin that modulate body weight by regulating  
XX energy utilisation - especially useful for treatment of obesity and  
XX diabetes.  
XX  
XX Claim 4; Page 1; 19pp; English.

CC The invention relates to specifically claimed peptides AAW45577-W45586 or  
CC their derivatives, analogues and variants, that modulate, specifically  
CC reduce, body weight, mainly by affecting energy utilisation. Also new  
CC are: (1) nucleic acid that encodes the peptides; (2) vectors containing  
CC the nucleic acid; and (3) host cells transformed with this vector. The  
CC peptides are used to treat nutritional or metabolic disorders,  
CC particularly obesity and diabetes.

CC Sequence 10 AA;  
SQ

Query Match 55.3%; Score 26; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DLHLV 10  
| | | | |  
| | | | |  
Db 1 DLHLV 5

RESULT 6  
ADK24124  
ID ADK24124 standard; peptide; 10 AA.  
XX



XX autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;  
 XX edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;  
 XX rheumatoid arthritis; chryoiditis; diabetic retinopathy;  
 XX transplant rejection; macular degeneration; neovascular glaucoma;  
 XX hemangioma; angiofibroma.  
 OS Homo sapiens.  
 XX WO2003086450-A1.  
 XX PD 23-OCT-2003.  
 XX PF 11-APR-2003; 2003WO-CU000004.  
 XX PR 15-APR-2002; 2002CU-00000076.  
 XX (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.  
 XX Bequet Romero M, Acevedo Castro BE, Gavilondo Cowley JV,  
 PI Fernandez Molina LE, Lopez Ochoa O, Silva Rodriguez RDC,  
 PI Musachio Lasa A, Galban Rodriguez E, Vazquez Blomquist DM;  
 XX WPI; 2003-833615/77.  
 XX PT Active immunization against angiogenic proteins, useful for treating e.g.  
 PT tumors and inflammation, particularly contains vascular endothelial  
 PT growth factor or its receptor.  
 PS Disclosure; Page 18; 53pp; Spanish.  
 XX The invention relates to an active immunization against angiogenic  
 CC proteins comprising administration of a vaccination composition (A),  
 CC optionally containing an adjuvant, that comprises polypeptides (I),  
 CC directly associated with an increase in angiogenesis, their variants, or  
 CC their encoding polynucleotides (II). Angiogenesis-associated polypeptides  
 CC are: members of the vascular endothelial growth factor (VEGF) family,  
 CC especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of  
 CC VEGF-B, or VEGF-C or -D; a (co-)receptor of VEGF, particularly VEGFR-1,  
 CC 2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental  
 CC growth factor. (A) is used for treatment or prevention of tumors in  
 CC mammals, particularly humans but also farm animals and pets, also many  
 CC other conditions associated with excessive angiogenesis, specifically  
 CC malignant or benign neoplasias (and their metastases), acute or chronic  
 CC inflammation, autoimmunity and eye diseases (claimed). Among the diseases  
 CC that may be treated are arthritis, endometriosis, arteriosclerosis,  
 CC edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,  
 CC psoriasis, rheumatoid arthritis, chryoiditis, diabetic retinopathy,  
 CC transplant rejection, macular degeneration, neovascular glaucoma,  
 CC hemangioma and angiofibroma. The method destroys cells that are the  
 CC source of angiogenic proteins, rather than just neutralizing the activity  
 CC of such proteins (as in passive immunization). This sequence represents  
 CC an immunisation peptide of the invention derived from the VEGF proteins.  
 XX  
 XX Sequence 9 AA:  
 XX  
 XX Query Match 53.2%; Score 25; DB 7; Length 9;  
 XX Best Local Similarity 66.7%; Pred. No. 1.7e+06;  
 XX Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 SLAMDLIH 9  
 DB 1 SLALILYH 9  
 XX  
 XX RESULT 9  
 XX ID AAR65116 standard; peptide; 10 AA.  
 XX AC AAR65116;  
 XX XX AAR65116;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 09-OCT-1995 (first entry)  
 XX

DE PSA immunogenic peptide 166-175.  
 XX  
 XX PSA; immunogenic peptide 166-175; cytotoxic C cells; in vitro activation;  
 XX cancer; AIDS; bacterial infections; malaria; fungal infections;  
 XX tuberculosis; hepatitis.  
 XX  
 XX OS Homo sapiens.  
 XX XX WO9504817-A1.  
 XX PD 16-FEB-1995.  
 XX PF 01-AUG-1994; 94WO-US008672.  
 XX PR 06-AUG-1993; 93US-00103401.  
 XX (CYTE-) CYTEL CORP.  
 XX Cellis E, Kubo R, Serra H, Tsai V, Wentworth P,  
 PI WPI; 1995-090895/12.  
 XX In vitro activation of cytotoxic T cells for selected killing of target  
 PT cells - for treating e.g. cancer, AIDS, hepatitis etc. by incubating them  
 PT with antigen presenting cells loaded with appropriate immunogenic  
 PT peptide.  
 XX Example 3; Page 35; 53pp; English.  
 XX AAR65109-R65145 are immunogenic peptides, they are used in a new method  
 CC for the in vitro activation of cytotoxic T cells (CTC). This is achieved  
 CC by incubating the CTCs with antigen presenting cells loaded with an  
 CC appropriate immunogenic peptide (e.g. one of the above peptides). By  
 CC selecting the peptides used the following diseases and infections can be  
 CC treated; cancer, AIDS, hepatitis, other viral and bacterial infections,  
 CC malaria and tuberculosis. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 10 AA:  
 XX  
 XX Query Match 53.2%; Score 25; DB 2; Length 10;  
 XX Best Local Similarity 44.4%; Pred. No. 2.9e+02;  
 XX Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 OY 2 LAMDLIHV 10  
 DB 2 LQCVDLVHI 10  
 XX  
 XX RESULT 10  
 XX ID AAG88173 standard; peptide; 10 AA.  
 XX AC AAG88173;  
 XX DT 11-SEP-2001 (first entry)  
 XX DE Saccharomyces cerevisiae peptide; SEQ ID NO: 3122.  
 XX KW Saccharomyces cerevisiae; complementary peptide; peptide identification;  
 XX drug discovery; drug design.  
 XX OS Saccharomyces cerevisiae.  
 XX XX WO200142276-A1.  
 XX PD 14-JUN-2001.  
 XX PF 13-DEC-2000; 2000WO-GB004773.  
 XX PR 13-DEC-1999; 99GB-00028471.  
 XX (PROT-) PROTEOM LTD.  
 XX

PI Roberts GW, Heal JR;  
 XX WPI; 2001-367863/38.  
 XX  
 XX Identifying complementary peptides by analysis of protein and nucleotide  
 PT sequence databases, useful in drug design.  
 XX  
 XX Example 5; Page 460; 488pp; English.  
 XX  
 CC The invention relates to the identification of complementary peptides by  
 CC analysis of protein and nucleotide sequence databases from higher  
 CC eukaryotic genomes, excluding human and plants. The specific  
 CC complementary peptides interact with their relevant target proteins  
 CC encoded in the eukaryote genome. The peptides may be used as reagents and  
 CC drugs for drug discovery and as lead ligands for drug design and  
 CC development. The present sequence is a complementary peptide from  
 CC Saccharomyces cerevisiae  
 CC  
 SQ Sequence 10 AA;  
 Query Match 53.2%; Score 25; DB 4; Length 10;  
 Best Local Similarity 62.5%; Pred. No. 2.9e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 LAMLDLH 9  
 Db 1 LQMKELH 8  
 RESULT 11  
 ID AAG8172  
 XX AAG8172 standard; peptide; 10 AA.  
 XX  
 AC AAG8172;  
 XX  
 DT 11-SEP-2001 (first entry)  
 XX  
 DE Saccharomyces cerevisiae peptide, SEQ ID NO: 3121.  
 XX  
 KW Saccharomyces cerevisiae; complementary peptide; peptide identification;  
 KM drug discovery; drug design.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 PN WO200142276-A1.  
 XX  
 PD 14-JUN-2001.  
 XX  
 PF 13-DEC-2000; 2000WO-GB004773.  
 XX  
 PR 13-DEC-1999; 99GB-00029471.  
 XX  
 PA (PROT-) PROTEOM LTD.  
 XX  
 PI Roberts GW, Heal JR;  
 XX  
 DR WPI; 2001-367863/38.  
 XX  
 PT Identifying complementary peptides by analysis of protein and nucleotide  
 PT sequence databases, useful in drug design.  
 XX  
 XX Example 5; Page 460; 488pp; English.  
 XX  
 CC The invention relates to the identification of complementary peptides by  
 CC analysis of protein and nucleotide sequence databases from higher  
 CC eukaryotic genomes, excluding human and plants. The specific  
 CC complementary peptides interact with their relevant target proteins  
 CC encoded in the eukaryote genome. The peptides may be used as reagents and  
 CC drugs for drug discovery and as lead ligands for drug design and  
 CC development. The present sequence is a complementary peptide from  
 CC Saccharomyces cerevisiae  
 CC  
 SQ Sequence 10 AA;

Query Match 53.2%; Score 25; DB 4; Length 10;  
 Best Local Similarity 62.5%; Pred. No. 2.9e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 LAMLDLH 9  
 Db 1 LQMKELH 8  
 RESULT 12  
 ID ABR22055  
 XX ABR22055 standard; peptide; 9 AA.  
 XX  
 AC ABR22055;  
 XX  
 DT 19-MAY-2003 (first entry)  
 XX  
 DE Human cancer-related protein 185P2C9 HLA peptide #1890.  
 XX  
 KW Human, cytostatic; vaccine; cancer; immune response; HLA;  
 KM human leukocyte antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283921-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 10-APR-2002; 2002WO-US011654.  
 XX  
 PR 10-APR-2001; 2001US-0282739P.  
 XX  
 PR 10-APR-2001; 2001US-0283112P.  
 XX  
 PR 25-APR-2001; 2001US-0286630P.  
 XX  
 PA (AGEN-) AGENSYS INC.  
 XX  
 PI Jakobovits A, Chailita-Bid PM, Faris W, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 XX  
 DR WPI; 2003-075555/07.  
 XX  
 PT New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides, useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.  
 XX  
 PS Claim 13; Page 341; 1021pp; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention  
 CC  
 SQ Sequence 9 AA;  
 Query Match 51.1%; Score 24; DB 6; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 SLAMDLH 9  
 Db 1 SLAMQNTLH 9



## RESULT 13

ABU03262 standard; protein; 9 AA.

ABU03262;

29-JAN-2003 (first entry)

Human expressed protein tag (EPT) #42.

Translational profiling; expressed protein tag; EPT; kinase; phosphatase; protease; protease inhibitor; transporter; cytoskeletal protein; receptor; transcription factor; cancer; MHC; major histocompatibility complex; myeloma; colon cancer; gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

Homo sapiens.

WO200278524-A2.

10-OCT-2002.

28-MAR-2002; 2002WO-US009671.

28-MAR-2001; 2001US-0279495P.

21-MAY-2001; 2001US-0292544P.

08-AUG-2001; 2001US-0310801P.

01-OCT-2001; 2001US-0326370P.

04-DEC-2001; 2001US-0356780P.

20-FEB-2002; 2002US-0358985P.

(ZYCO-) ZYCO INC.

Chicz RM, Tomlinson AJ, Urban RG;

WPI; 2003-040607/03.

Claim 10; SEQ ID NO 42; 134pp; English.

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at [ftp://wipo.int/publ/published\\_pct\\_sequences](http://wipo.int/publ/published_pct_sequences)

Sequence 9 AA;

Query Match 51.1%; Score 24; DB 6; Length 9;

Best Local Similarity 50.0%; Pred. No. 1.7e+06; Mismatches 0; Indels 0; Gaps 0;

Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

4 MDDLIH 9

2 MVDIIH 7

## RESULT 14

AAU94381 standard; peptide; 10 AA.

AAU94381;

02-JUL-2002 (first entry)

Human novel protein CatrF2E11 HLA binding peptide #164.

Human; human leukocyte antigen; HLA; immunogen; 83P2H3; CatrF2E11; calcium transport protein; cancer; prostate cancer; cytostatic; chromosome 7q34; chromosome 12q24.1; T cell; B cell.

Homo sapiens.

WO200214361-A2.

21-FEB-2002.

17-AUG-2001; 2001WO-US025782.

17-AUG-2000; 2000US-0226329P.

(AGEN-) AGENSYS INC.

Rattano AB, Chalitta-Elid PM, Faris M, Saffran DC, Afar DEH;

Levin E, Hubert RS, Ge W, Jakobovits A;

WPI; 2002-269179/31.

Monitoring 83P2H3 gene products for monitoring the presence of cancer in a subject, comprises determining the status of 83P2H3 gene products in a tissue sample from the subject and comparing it to a normal sample.

Example 11; Page 176; 270pp; English.

The invention relates to monitoring 83P2H3 (a calcium transport protein whose gene is located on chromosome 7q34) gene products in a biological sample from a patient who has or is suspected of having cancer (especially prostate cancer), comprising: (a) determining the status of 83P2H3 gene products expressed by cells in a tissue sample from an individual and (b) comparing the status to the status of 83P2H3 gene products in a normal sample. Also included are modulators of 83P2H3 function or status, generating antibodies/immune response against 83P2H3 (or related protein CatrF2E11 whose gene is located on chromosome 12q24.1) using identified HLA (human leukocyte antigen) binding peptides derived from the protein, delivering a cytotoxic agent to a cell expressing 83P2H3 by conjugating the agent to an anti-83P2H3 antibody, a recombinant protein comprising an antigen-binding region of the antibody, a non-human transgenic animal that produces the recombinant protein, a hybridoma that produces the recombinant protein, a single-chain monoclonal antibody that comprises the variable domains of the heavy and light chains of the anti-83P2H3 antibody, a vector comprising a polynucleotide that encodes the monoclonal antibody and inducing an immune response to a 83P2H3 protein, by providing a 83P2H3-related protein that comprises a T cell or B cell epitope, and contacting the epitope with an immune system T cell or B cell, respectively. The method is useful for monitoring 83P2H3 gene products in a biological sample for monitoring the presence of cancer in an individual. The modulator is useful for inhibiting the growth of cancer cells that express 83P2H3, for treating cancer and the vector is useful for treating a patient with a cancer that expresses 83P2H3. The immunological methods are useful for generating an immune response against 83P2H3, and for detecting the presence of 83P2H3-related protein or polynucleotide in a biological sample from a patient who has or is suspected of having cancer. The antibody is useful in prostate cancer diagnosis, prognosis, imaging methodologies and treatment, to detect and quantify 83P2H3 and mutant 83P2H3-related proteins, for purifying a 83P2H3-related protein, for isolating 83P2H3 homologues/related molecules, and for generating anti-idiotypic antibodies that mimic the 83P2H3 protein. The present sequence is an HLA binding peptide motif from 83P2H3 or its related protein CatrF2E11

XX Sequence 10 AA;

Query Match 51.1%; Score 24; DB 5; Length 10;  
Best Local Similarity 66.7%; Pred. No. 4.5e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LAMLDLHV 10  
|||  
2 LAALLLVHV 10

OY 1 SLAMLDLHV 10  
|||  
1 SLGLSLVAV 10  
Search completed: November 30, 2004, 08:09:32  
Job time : 57.333 secs

## RESULT 15

ADK24903  
ID ADK24903 standard; peptide; 10 AA.

AC ADK24903;

DT 06-MAY-2004 (first entry)

DE Human 98P4B6 derived motif-bearing CTL peptide epitope #1013.

XX human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney; prostate;  
XX gene therapy; genetic abnormality; transgenic; knockout animal;  
XX cytosolic; epitope.

OS Homo sapiens.

PN W02003087306-A2.

PD 23-OCT-2003.

PF 04-APR-2003; 2003WO-US010462.

PR 05-APR-2002; 2002US-0370387P.

PR 06-JUN-2002; 2002US-00165044.

PR 20-DEC-2002; 2002US-0435480P.

PA (AGEN-) AGENSYS INC.  
XX Chailita-Eid PM, Raitano AB, Paris M, Ge W, Jakobovits A;

DR WPI; 2003-903158/82.

PT A composition comprising 98P4B6 proteins and nucleic acid molecules for  
XX detecting, preventing, prognosing and/or treating cancers that express  
XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

PS Claim 1; Page 152; 616pp; English.

XX This invention relates to a novel human gene 98P4B6 (also known as STEAP-  
CC 2, six transmembrane epithelial antigen of prostate-2) located on  
CC chromosome 7q21 and the encoded protein and variants derived thereof.  
CC Specifically, it refers to the expression pattern of this gene in adult  
CC normal tissues and its aberrant over-expression in various cancers  
CC including breast, colon, lung, kidney and prostate. The present invention  
CC describes compositions and methods useful for detecting, preventing,  
CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the  
CC polynucleotide can be used for gene therapy purposes, for monitoring  
CC genetic abnormalities and for generating transgenic or 'knockout'  
CC animals, which can be useful for the development and screening of  
CC therapeutically useful reagents. The encoded proteins may also be used in  
CC generating antibodies and vaccines, as well as in identifying ligands and  
CC cellular constituents that bind to 98P4B6 gene products. Accordingly,  
CC these compositions exhibit cytostatic activities. This peptide sequence  
CC is a motif bearing CTL peptide epitope derived from a human 98P4B6  
CC protein variant of the invention.

XX Sequence 10 AA;

Query Match 51.1%; Score 24; DB 7; Length 10;  
Best Local Similarity 60.0%; Pred. No. 4.5e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 08:13:53 ; Search time 44 Seconds  
(without alignments)

80.612 Million cell updates/sec

Title: US-10-008-377a-1

Perfect score: 47

Sequence: 1 SLAMDILHV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 177356

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PTCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/PTCTUS\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PTCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10D\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US10D\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	10	US-10-008-377-1	Sequence 1, Appli
2	43	91.5	9	US-10-008-377-3	Sequence 3, Appli
3	39	83.0	9	US-10-008-377-4	Sequence 4, Appli
4	26	55.3	10	US-09-844-774-9	Sequence 9, Appli
5	26	55.3	10	US-10-290-144-9	Sequence 9, Appli
6	24	51.1	8	US-10-064-903-2	Sequence 2, Appli
7	24	51.1	10	US-09-932-165-364	Sequence 364, App
8	22.5	47.9	10	US-10-024-652-1548	Sequence 1548, Ap
9	22.5	47.9	10	US-10-024-652-1816	Sequence 1816, Ap
10	22	46.8	7	US-09-758-128-30	Sequence 30, Appli
11	22	46.8	7	US-09-758-426-30	Sequence 30, Appli
12	22	46.8	7	US-09-758-198-30	Sequence 30, Appli
13	22	46.8	7	US-09-861-661-30	Sequence 30, Appli

14	22	46.8	7	US-10-400-991-27	Sequence 27, Appli
15	22	46.8	8	US-10-367-580-170	Sequence 170, App
16	22	46.8	8	US-10-367-593-170	Sequence 170, App
17	22	46.8	8	US-10-367-594-170	Sequence 170, App
18	22	46.8	8	US-10-367-654-170	Sequence 170, App
19	22	46.8	8	US-10-367-658-170	Sequence 170, App
20	22	46.8	8	US-10-367-658-170	Sequence 170, App
21	22	46.8	8	US-10-367-674-170	Sequence 170, App
22	22	46.8	8	US-09-393-634-88	Sequence 88, Appli
23	22	46.8	9	US-09-753-126-119	Sequence 119, Appli
24	22	46.8	9	US-09-886-896A-83	Sequence 83, Appli
25	22	46.8	9	US-09-932-165-225	Sequence 225, App
26	22	46.8	9	US-09-932-165-441	Sequence 441, App
27	22	46.8	9	US-09-932-165-841	Sequence 841, App
28	22	46.8	9	US-09-932-165-968	Sequence 968, App
29	22	46.8	9	US-10-383-982-88	Sequence 88, Appli
30	22	46.8	9	US-10-334-726-157	Sequence 157, Appli
31	22	46.8	9	US-10-024-652-3	Sequence 3, Appli
32	22	46.8	9	US-10-024-652-105	Sequence 105, App
33	22	46.8	9	US-10-024-652-147	Sequence 147, App
34	22	46.8	9	US-10-024-652-136	Sequence 136, App
35	22	46.8	9	US-10-024-652-442	Sequence 442, App
36	22	46.8	9	US-10-024-652-510	Sequence 510, App
37	22	46.8	9	US-10-024-652-537	Sequence 537, App
38	22	46.8	9	US-10-024-652-627	Sequence 627, App
39	22	46.8	9	US-10-024-652-642	Sequence 642, App
40	22	46.8	9	US-10-024-652-888	Sequence 888, App
41	22	46.8	9	US-10-024-652-840	Sequence 840, App
42	22	46.8	9	US-10-024-652-1046	Sequence 1046, App
43	22	46.8	9	US-10-024-652-1110	Sequence 1110, App
44	22	46.8	9	US-10-024-652-1111	Sequence 1111, App
45	22	46.8	9	US-10-024-652-1201	Sequence 1201, App

## ALIGNMENTS

RESULT 1  
US-10-008-377-1  
Sequence 1, Application US/10008377  
Publication No. US20030157101A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
TITLE OF INVENTION: Immunogenic ALK Peptides  
FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008,377  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-1

Query Match 100.0% Score 47; DB 14; Length 10;  
Best Local Similarity 100.0% Pred. No. 0.066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;  
QY 1 SLAMDILHV 10  
DB 1 SLAMDILHV 10  
RESULT 2  
US-10-008-377-3  
Sequence 3, Application US/10008377  
Publication No. US20030157101A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
APPLICANT: Passoni, Lorena  
TITLE OF INVENTION: Immunogenic ALK Peptides

FILE REFERENCE: 045922/241203  
 CURRENT APPLICATION NUMBER: US/10/008,377  
 CURRENT FILING DATE: 2001-11-15  
 NUMBER OF SEQ ID NOS: 7  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 3  
 LENGTH: 9  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-008-377-3

Query Match  
 Best Local Similarity 91.5%; Score 43; DB 14; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAMIDLHV 10  
 DB 1 LAMIDLHV 9

RESULT 3  
 US-10-008-377-4  
 Sequence 4, Application US/10008377  
 Publication No. US20030157101A1  
 GENERAL INFORMATION:  
 APPLICANT: Gambacorti-Passerini, Carlo  
 APPLICANT: Passoni, Lorena  
 TITLE OF INVENTION: Immunogenic ALK Peptides  
 FILE REFERENCE: 045922/241203  
 CURRENT APPLICATION NUMBER: US/10/008,377  
 CURRENT FILING DATE: 2001-11-15  
 NUMBER OF SEQ ID NOS: 7  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 4  
 LENGTH: 9  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-008-377-4

Query Match  
 Best Local Similarity 83.0%; Score 39; DB 14; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMIDLHV 10  
 DB 1 AMIDLHV 8

RESULT 4  
 US-09-844-774-9  
 Sequence 9, Application US/09844774  
 Publication No. US20020037553A1  
 GENERAL INFORMATION:  
 APPLICANT: Al-Barazangi, Kamal A.  
 APPLICANT: Arch, Jonathan Robert Sanders  
 APPLICANT: Camilleri, Patrick  
 APPLICANT: Neville, William Arthur  
 TITLE OF INVENTION: Fragments of Leptin (OB protein)  
 FILE REFERENCE: P31505-C1  
 CURRENT APPLICATION NUMBER: US/09/844,774  
 CURRENT FILING DATE: 2001-04-27  
 PRIOR APPLICATION NUMBER: 09/154,866  
 PRIOR FILING DATE: 1998-12-04  
 PRIOR APPLICATION NUMBER: PCT/EP97/02968  
 PRIOR FILING DATE: 1997-06-04  
 PRIOR APPLICATION NUMBER: GB 9703493.8  
 PRIOR FILING DATE: 1997-02-20  
 PRIOR APPLICATION NUMBER: GB 9618540.0  
 PRIOR FILING DATE: 1996-09-05  
 PRIOR APPLICATION NUMBER: GB 9611775.9  
 PRIOR FILING DATE: 1996-06-06  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 9  
 LENGTH: 10  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-844-774-9

Query Match  
 Best Local Similarity 55.3%; Score 26; DB 9; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DLHV 10  
 DB 1 DLHV 5

RESULT 5  
 US-10-290-144-9  
 Sequence 9, Application US/10290144  
 Publication No. US20030092126A1  
 GENERAL INFORMATION:  
 APPLICANT: Al-Barazangi, Kamal A.  
 APPLICANT: Arch, Jonathan Robert Sanders  
 APPLICANT: Camilleri, Patrick  
 APPLICANT: Neville, William Arthur  
 TITLE OF INVENTION: Fragments of Leptin (OB protein)  
 FILE REFERENCE: P31505C2  
 CURRENT APPLICATION NUMBER: US/10/290,144  
 CURRENT FILING DATE: 2002-11-07  
 PRIOR APPLICATION NUMBER: 09/844,774  
 PRIOR FILING DATE: 2001-04-27  
 PRIOR APPLICATION NUMBER: 09/194,866  
 PRIOR FILING DATE: 1998-12-04  
 PRIOR APPLICATION NUMBER: PCT/EP97/02968  
 PRIOR FILING DATE: 1997-06-04  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 9  
 LENGTH: 10  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-290-144-9

Query Match  
 Best Local Similarity 55.3%; Score 26; DB 14; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DLHV 10  
 DB 1 DLHV 5

RESULT 6  
 US-10-064-903-2  
 Sequence 2, Application US/10064903  
 Publication No. US20030059912A1  
 GENERAL INFORMATION:  
 APPLICANT: Biotechnische Gesellschaft fur biotechnologische Entwicklung und Consulting  
 APPLICANT: MBH  
 TITLE OF INVENTION: HYBRID PROTEIN FOR INHIBITING THE DEGRADATION OF MASTOCYTES AND  
 FILE REFERENCE: BIO-001PCT-CIP  
 CURRENT APPLICATION NUMBER: US/10/064,903  
 CURRENT FILING DATE: 2002-08-27  
 PRIOR APPLICATION NUMBER: US 09/700,540  
 PRIOR FILING DATE: 2001-01-19  
 NUMBER OF SEQ ID NOS: 2  
 SOFTWARE: Patent version 3.1  
 SEQ ID NO 2  
 LENGTH: 8  
 TYPE: PRT  
 ORGANISM: Clostridium tetrani  
 US-10-064-903-2

Query Match 51.1%; Score 24; DB 14; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 DLIHV 10  
DB 2 DLIHV 6

RESULT 7  
US-09-932-165-364  
; Sequence 364, Application US/09932165  
; Publication No. US20030134784A1  
; GENERAL INFORMATION:  
; APPLICANT: RAITANO, ARTHUR  
; APPLICANT: CHALLITA-EID, PIA M.  
; APPLICANT: FARIS, MARY  
; APPLICANT: SAFRAN, DOUGLAS  
; APPLICANT: AFAR, DANIEL  
; APPLICANT: LEVIN, ELANA  
; APPLICANT: HUBERT, RENE  
; APPLICANT: GE, WANGMAO  
; APPLICANT: JAKOBYTS, AVA  
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
; TITLE OF INVENTION: 83P2H3 AND CATF2E11 USEFUL IN TREATMENT AND  
; FILE REFERENCE: 51158-20014.00  
; CURRENT APPLICATION NUMBER: US/09/932,165  
; CURRENT FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/226,329  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 1508  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 364  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-364

Query Match 51.1%; Score 24; DB 10; Length 10;  
Best Local Similarity 66.7%; Pred. No. 7.6e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
DB 2 LAALLVHV 10

RESULT 8  
US-10-024-652-1548  
; Sequence 1548, Application US/10024652  
; Publication No. US20030219738A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc  
; APPLICANT: Challita-Eid, Pia M.  
; APPLICANT: Faris, Mary  
; APPLICANT: Afar, Daniel E.H.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Mitchell, Steve Chappell  
; APPLICANT: Levin, Elana  
; APPLICANT: Morrison, Karen Jane Meyrick  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Jakobovits, Ava  
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and  
; FILE REFERENCE: 51158-20025.00  
; CURRENT APPLICATION NUMBER: US/10/024,652  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/256,210  
; PRIOR FILING DATE: 2000-12-15

; NUMBER OF SEQ ID NOS: 2598  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1548  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: homo sapien

US-10-024-652-1548

Query Match 47.9%; Score 22.5; DB 14; Length 10;  
Best Local Similarity 70.0%; Pred. No. 1.4e+03;  
Matches 7; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 SLA-MIDLH 9  
DB 1 SLAINTDALH 10

RESULT 9  
US-10-024-652-1816  
; Sequence 1816, Application US/10024652  
; Publication No. US20030219738A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc  
; APPLICANT: Challita-Eid, Pia M.  
; APPLICANT: Faris, Mary  
; APPLICANT: Afar, Daniel E.H.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Mitchell, Steve Chappell  
; APPLICANT: Levin, Elana  
; APPLICANT: Morrison, Karen Jane Meyrick  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Jakobovits, Ava  
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and  
; FILE REFERENCE: 51158-20025.00  
; CURRENT APPLICATION NUMBER: US/10/024,652  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/256,210  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 2598  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1816  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: homo sapien

Query Match 47.9%; Score 22.5; DB 14; Length 10;  
Best Local Similarity 70.0%; Pred. No. 1.4e+03;  
Matches 7; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 SLA-MIDLH 9  
DB 1 SLAINTDALH 10

RESULT 10  
US-09-758-128-30  
; Sequence 30, Application US/09758128  
; Patent No. US20020107187A1  
; GENERAL INFORMATION:  
; APPLICANT: KINGSTON, David J.  
; APPLICANT: GERARDY, No. US20020107187A1man L.  
; APPLICANT: WESTBROOK, Simon L.  
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS  
; TITLE OF INVENTION: PEPTIDES, ANTIBODIES, VACCINES AND USES THEREOF  
; FILE REFERENCE: 016786/0214  
; CURRENT APPLICATION NUMBER: US/09/758,128  
; CURRENT FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: 09/194,218  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PN9390

;; PRIOR FILING DATE: 1996-05-22  
;; NUMBER OF SEQ ID NOS: 58  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO: 30  
;; LENGTH: 7  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-758-128-30

Query Match 46.8%; Score 22; DB 9; Length 7;  
Best Local Similarity 60.0%; Pred. No. 1.4e+06;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LDLLH 9  
:|:|  
Db 1 MDMLH 5

RESULT 11  
US-09-758-426-30  
; Sequence 30, Application US/09758426  
; Patent No. US20020169116A1  
; GENERAL INFORMATION:  
; APPLICANT: KINGSTON, David J.  
; APPLICANT: GERRATY, No. US20020169116A1man L.  
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS  
; FILE REFERENCE: 016786/0214  
; CURRENT APPLICATION NUMBER: US/09/758,426  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: 09/194,218  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PN9990  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 30  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-426-30

Query Match 46.8%; Score 22; DB 9; Length 7;  
Best Local Similarity 60.0%; Pred. No. 1.4e+06;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LDLLH 9  
:|:|  
Db 1 MDMLH 5

RESULT 12  
US-09-758-198-30  
; Sequence 30, Application US/09758198  
; Publication No. US20020187925A1  
; GENERAL INFORMATION:  
; APPLICANT: KINGSTON, David J.  
; APPLICANT: GERRATY, No. US20020187925A1man L.  
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS  
; FILE REFERENCE: 016786/0214  
; CURRENT APPLICATION NUMBER: US/09/758,198  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-02-05  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-22  
; PRIOR APPLICATION NUMBER: PCT/AU97/00312  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-05-22  
; PRIOR APPLICATION NUMBER: AU PN9990  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO: 30  
;; LENGTH: 7  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-758-198-30

Query Match 46.8%; Score 22; DB 9; Length 7;  
Best Local Similarity 60.0%; Pred. No. 1.4e+06;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LDLLH 9  
:|:|  
Db 1 MDMLH 5

RESULT 13  
US-09-861-661-30  
; Sequence 30, Application US/09861661  
; Publication No. US20030045676A1  
; GENERAL INFORMATION:  
; APPLICANT: KINGSTON, DAVID J.  
; APPLICANT: WESTBROOK, SIMON L.  
; TITLE OF INVENTION: PEPTIDES, ANTIBODIES, VACCINES & USES THEREOF  
; FILE REFERENCE: 054270/0135  
; CURRENT APPLICATION NUMBER: US/09/861,661  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: 09/194,218  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PN9990  
; PRIOR FILING DATE: 1996-05-22  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 30  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-861-661-30

Query Match 46.8%; Score 22; DB 10; Length 7;  
Best Local Similarity 60.0%; Pred. No. 1.4e+06;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LDLLH 9  
:|:|  
Db 1 MDMLH 5

RESULT 14  
US-10-400-991-27  
; Sequence 27, Application US/10400991  
; Publication No. US20030224417A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Glucksmann, Maria A.  
; APPLICANT: Welch, Nadine S.  
; APPLICANT: Hunter, John J.  
; APPLICANT: White, David  
; APPLICANT: Macbeth, Kyle J.  
; APPLICANT: Williamson, Mark J.  
; APPLICANT: Meyers, Rachel E.  
; TITLE OF INVENTION: 14400, 2838, 14618, 14274, 32164,  
; TITLE OF INVENTION: 39404, 38911, 26904, 31237, 18057, 16405, 32705, 23224,  
; TITLE OF INVENTION: 27423, 32700, 32712 AND 12216, NOVEL SEVEN-TRANSMEMBRANE  
; FILE REFERENCE: ME103-0240NM1  
; CURRENT APPLICATION NUMBER: US/10/400,991  
; PRIOR FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: US 10/190,469  
; PRIOR FILING DATE: 2002-07-05  
; PRIOR APPLICATION NUMBER: US 09/439,159  
; PRIOR FILING DATE: 1999-11-12

PRIOR APPLICATION NUMBER: US 09/137,063  
PRIOR FILING DATE: 1998-08-20  
PRIOR APPLICATION NUMBER: US 10/167,192  
PRIOR FILING DATE: 2003-06-11  
PRIOR APPLICATION NUMBER: US 09/420,187  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: US 09/173,869  
PRIOR FILING DATE: 1998-10-16  
PRIOR APPLICATION NUMBER: US 10/339,056  
PRIOR FILING DATE: 2003-01-09  
PRIOR APPLICATION NUMBER: US 09/377,429  
PRIOR FILING DATE: 1999-08-19  
PRIOR APPLICATION NUMBER: US 09/136,726  
PRIOR FILING DATE: 1998-08-18  
PRIOR APPLICATION NUMBER: US 09/911,583  
PRIOR FILING DATE: 2001-07-24  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 83  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 27  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Amino Acid Fragment  
US-10-400-991-27

Query Match 46.8%; Score 22; DB 14; Length 7;  
Best Local Similarity 71.4%; Pred. No. 1.4e+06;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LAMDDL 8  
DB 1 LAVADLL 7

RESULT 15  
US-10-367-580-170  
Sequence 170, Application US/10367580  
Publication No. US20040071720A1  
GENERAL INFORMATION:  
APPLICANT: Rothman, James E.  
APPLICANT: Hartl, F. Ulrich  
APPLICANT: Hoe, Mee H.  
APPLICANT: Houghton, Alan  
APPLICANT: Takeuchi, Yoshizumi  
APPLICANT: Mayhew, Mark  
TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies  
FILE REFERENCE: 11746/461061  
CURRENT APPLICATION NUMBER: US/10/367,580  
CURRENT FILING DATE: 2003-02-14  
PRIOR APPLICATION NUMBER: US 09/794,832  
PRIOR FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: US 09/011,645  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: PCT/US96/13363  
PRIOR FILING DATE: 1996-08-16  
PRIOR APPLICATION NUMBER: US 60/002,490  
PRIOR FILING DATE: 1995-08-18  
PRIOR APPLICATION NUMBER: US 60/002,479  
PRIOR FILING DATE: 1995-08-18  
NUMBER OF SEQ ID NOS: 349  
SOFTWARE: WordPerfect 8.0 for Windows  
SEQ ID NO 170  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic peptide  
US-10-367-580-170

Query Match 46.8%; Score 22; DB 15; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 6 DILL 9  
DB 3 DILL 6

Search completed: November 30, 2004, 08:31:00  
Job time : 45 secs

**This Page Blank (uspio)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: November 30, 2004, 07:41:57 ; Search time 27 Seconds

(without alignments)  
24.562 Million cell updates/sec

Title: US-10-008-377A-1

Perfect score: 47

Sequence: 1 SLAMDLHLV 10

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 110780

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

1: /cgn2\_6/ptodata/1/aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/aa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	53.2	10	3	US-08-159-339A-649
2	23	48.9	10	4	US-09-601-729-268
3	22	46.8	9	4	US-09-393-634-88
4	21	44.7	7	3	US-09-258-754-114
5	21	44.7	7	3	US-09-047-107-414
6	21	44.7	7	4	US-09-722-250D-414
7	21	44.7	7	4	US-09-676-45A-414
8	21	44.7	8	2	US-08-917-320-13
9	21	44.7	8	5	PCT-US95-04611A-13
10	21	44.7	9	1	US-08-217-188A-3
11	21	44.7	9	1	US-08-687-226-3
12	21	44.7	9	3	US-08-687-725B-3
13	21	44.7	9	3	US-09-007-748-3
14	21	44.7	9	3	US-09-053-003-16
15	21	44.7	9	3	US-09-183-931-31
16	21	44.7	9	3	US-09-705-160-31
17	21	44.7	9	4	US-08-197-484-81
18	21	44.7	9	4	US-08-197-484-140
19	21	44.7	9	4	US-09-818-094-16
20	21	44.7	9	5	PCT-US95-02121-81
21	21	44.7	9	5	PCT-US95-02121-140
22	21	44.7	10	1	US-08-217-188A-20
23	21	44.7	10	1	US-08-687-226-20
24	21	44.7	10	3	US-08-687-725B-20
25	21	44.7	10	3	US-09-007-748-20
26	21	44.7	10	4	US-08-197-484-82
27	21	44.7	10	4	US-08-197-484-141

28	21	44.7	10	4	US-08-980-326-18	Sequence 18, App1
29	21	44.7	10	5	PCT-US95-02121-82	Sequence 82, App1
30	21	44.7	10	5	PCT-US95-02121-141	Sequence 141, App1
31	20	42.6	5	4	US-09-082-358B-10	Sequence 10, App1
32	20	42.6	7	1	US-08-166-930-16	Sequence 16, App1
33	20	42.6	7	2	US-08-727-045A-16	Sequence 16, App1
34	20	42.6	7	3	US-09-223-587-1	Sequence 1, App1
35	20	42.6	7	3	US-09-312-314C-1	Sequence 1, App1
36	20	42.6	7	3	US-09-312-314C-1	Sequence 1, App1
37	20	42.6	7	4	US-09-516-488A-8	Sequence 8, App1
38	20	42.6	7	4	US-09-408-172-16	Sequence 16, App1
39	20	42.6	7	4	US-09-394-234-8	Sequence 8, App1
40	20	42.6	9	3	US-08-817-547A-2	Sequence 2, App1
41	20	42.6	9	3	US-09-008-481A-13	Sequence 13, App1
42	20	42.6	9	3	US-09-309-592-13	Sequence 13, App1
43	20	42.6	10	1	US-08-843-035-17	Sequence 17, App1
44	20	42.6	10	1	US-08-843-035-18	Sequence 18, App1
45	20	42.6	10	1	US-08-843-035-19	Sequence 19, App1

#### ALIGNMENTS

RESULT 1  
US-08-159-339A-649  
Sequence 649, Application US/08159339A  
Patent No. 6037135  
GENERAL INFORMATION:  
APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Esben  
TITLE OF INVENTION: HLA Binding peptides and their  
NUMBER OF SEQUENCES: 1254  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/159,339A  
FILING DATE: 29-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,666  
FILING DATE: 07-AUG-1992  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen Lawver  
REGISTRATION NUMBER: 32762  
REFERENCE/DOCKET NUMBER: 018623-005030US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX:  
INFORMATION FOR SEQ ID NO: 649:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-159-339A-649

Query Match 53.2%; Score 25; DB 3; Length 10;  
 Best Local Similarity 44.4%; Pred. No. 1.6e+02;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAMDLHV 10  
 : : : : :  
 DB 2 LQCVDLVHI 10

RESULT 2

US-09-601-729-268  
 ; Sequence 268, Application US/09601729  
 ; Patent No. 6683052  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THIAM, KADER  
 ; APPLICANT: AURIAULT, CLAUDE  
 ; APPLICANT: GRAS-MASSÉ, HELENE  
 ; APPLICANT: LOTING, ESTELLE  
 ; APPLICANT: VERMAERDE, CLAUDE  
 ; APPLICANT: GUILLET, JEAN GERARD  
 ; TITLE OF INVENTION: LIPOPEPTIDES CONTAINING AN INTERFERON FRAGMENT AND USES  
 ; FILE REFERENCE: US-97-AU-IN  
 ; CURRENT APPLICATION NUMBER: US/09/601,729  
 ; PRIOR FILING DATE: 2000-11-20/601,729  
 ; PRIOR APPLICATION NUMBER: PCT/FR99/00259  
 ; PRIOR FILING DATE: 1999-02-05  
 ; PRIOR APPLICATION NUMBER: 98 01439  
 ; PRIOR FILING DATE: 1998-02-06  
 ; NUMBER OF SEQ ID NOS: 281  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 268  
 ; LENGTH: 10  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 ; OTHER INFORMATION: peptide  
 US-09-601-729-268

Query Match 48.9%; Score 23; DB 4; Length 10;  
 Best Local Similarity 71.4%; Pred. No. 3.5e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AMLDLH 9  
 : : : : :  
 DB 2 AALDLH 8

RESULT 3

US-09-393-634-88  
 ; Sequence 88, Application US/09393634  
 ; Patent No. 6558910  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Zuker, Charles S.  
 ; APPLICANT: Adler, Jon Elliot  
 ; APPLICANT: Ryba, Nick  
 ; APPLICANT: Mueller, Ken  
 ; APPLICANT: Hoon, Mark  
 ; APPLICANT: The Regents of the University of California  
 ; APPLICANT: The Government of the United States of America  
 ; APPLICANT: as represented by the Secretary of the  
 ; APPLICANT: Department of Health and Human Services  
 ; TITLE OF INVENTION: SF, a No. 6558910e1 Family of Taste Receptors  
 ; FILE REFERENCE: 023075-098000US  
 ; CURRENT APPLICATION NUMBER: US/09/393,634  
 ; CURRENT FILING DATE: 1999-09-10  
 ; NUMBER OF SEQ ID NOS: 92  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 88  
 ; LENGTH: 9

TYPE: PRT

ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: SF01 signature  
 ; OTHER INFORMATION: sequence 1, amino acids encoded by PCR primers  
 ; OTHER INFORMATION: identifying polymorphic variants, interspecies  
 ; OTHER INFORMATION: homologs and alleles of Sf family members  
 US-09-393-634-88

Query Match 46.8%; Score 22; DB 4; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 3.8e+05;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMDL 8  
 : : : : :  
 DB 2 MAPDL 8

RESULT 4

US-09-258-754-414  
 ; Sequence 414, Application US/09258754  
 ; Patent No. 6174687  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruostandi, Erkki  
 ; APPLICANT: Pasqualini, Renata  
 ; APPLICANT: Rajotte, Daniel  
 ; TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using  
 ; FILE REFERENCE: P-LJ 3443  
 ; CURRENT APPLICATION NUMBER: US/09/258,754  
 ; CURRENT FILING DATE: 1999-02-26  
 ; EARLIER APPLICATION NUMBER: 09/042,107  
 ; EARLIER FILING DATE: 1998-03-13  
 ; NUMBER OF SEQ ID NOS: 452  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO 414  
 ; LENGTH: 7  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 US-09-258-754-414

Query Match 44.7%; Score 21; DB 3; Length 7;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMD 6  
 : : : : :  
 DB 1 SLAMD 6

RESULT 5

US-09-042-107-414  
 ; Sequence 414, Application US/09042107  
 ; Patent No. 6232287  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruostandi, Erkki  
 ; APPLICANT: Pasqualini, Renata  
 ; APPLICANT: Rajotte, Daniel  
 ; TITLE OF INVENTION: Molecules that Home to Various Selected Organs or  
 ; FILE REFERENCE: P-LJ 2892  
 ; CURRENT APPLICATION NUMBER: US/09/042,107  
 ; CURRENT FILING DATE: 1998-03-13  
 ; NUMBER OF SEQ ID NOS: 436  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO 414  
 ; LENGTH: 7  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 US-09-042-107-414

Tue Nov 30 08:49:15 2004

us-10-008-377a-1.closed.ral

Page 3

Query Match 44.7%; Score 21; DB 3; Length 7;  
Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMD 6  
DB 1 SLAMD 6

RESULT 6  
US-09-722-250D-414  
Sequence 414, Application US/09722250D  
Patent No. 6610651  
GENERAL INFORMATION:  
APPLICANT: Rucsiabhi, Erkki  
APPLICANT: Pasqualini, Renata  
TITLE OF INVENTION: Molecules that Home to Various Selected Organs or  
FILE REFERENCE: P-LJ 4514  
CURRENT APPLICATION NUMBER: US/09/722,250D  
CURRENT FILING DATE: 2000-11-22  
PRIOR APPLICATION NUMBER: US 09/042,107  
NUMBER OF SEQ ID NOS: 437  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 414  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-722-250D-414

Query Match 44.7%; Score 21; DB 4; Length 7;  
Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMD 6  
DB 1 SLAMD 6

RESULT 7  
US-09-676-475A-414  
Sequence 414, Application US/09676475A  
Patent No. 6784153  
GENERAL INFORMATION:  
APPLICANT: Rucsiabhi, Erkki  
APPLICANT: Pasqualini, Renata  
APPLICANT: Rajotte, Daniel  
TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using  
FILE REFERENCE: P-LA 4377  
CURRENT APPLICATION NUMBER: US/09/676,475A  
CURRENT FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 09/042,107  
NUMBER OF SEQ ID NOS: 452  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 414  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-676-475A-414

Query Match 44.7%; Score 21; DB 4; Length 7;  
Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 SLAMD 6

RESULT 8  
US-08-917-320-13  
Sequence 13, Application US/08917320  
Patent No. 5824508  
GENERAL INFORMATION:  
APPLICANT: Spaete, Richard and Jackman, Winthrop, T.  
TITLE OF INVENTION: No. 5824508 Splicing Variants of gp350/220  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooley Godward Castro Huddleson & Tatum  
STREET: 5 Palo Alto Square  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/917,320  
FILING DATE: 25-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/229,291  
FILING DATE: April 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Luan Casert  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: AVIR-003/000US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-843-5163  
TELEFAX: 415-857-0663  
TELEX: 380816 CooleyPA  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-917-320-13

Query Match 44.7%; Score 21; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 3.8e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMD 6  
DB 1 SLAMD 6

RESULT 9  
PCT-US95-04611A-13  
Sequence 13, Application PC/TUS9504611A  
GENERAL INFORMATION:  
APPLICANT: Spaete, Richard and Jackman, Winthrop, T.  
TITLE OF INVENTION: Non Splicing Variants of gp350/220  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooley Godward Castro Huddleson & Tatum  
STREET: 5 Palo Alto Square  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04611A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/229,291
FILING DATE: April 18, 1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Luanh Csery
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: AVIR-003/00US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-843-5163
TELEFAX: 415-857-0663
TELEX: 380816 CooleyPA
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPES: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
PCT-US95-04611A-13

Query Match
Best Local Similarity 44.7%; Score 21; DB 5; Length 8;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

CY 1 SLAMD 6
Db 1 NISMLD 6

RESULT 10
US-08-217-188A-3
Sequence 3, Application US/08217188A
Patent No. 5534724
GENERAL INFORMATION:
APPLICANT: Melief, Cornelis J. M.
APPLICANT: Visseren, M. J. W.
APPLICANT: Kast, W. M.
APPLICANT: van der Bruggen, Pierre
APPLICANT: Boon-Falleur, Thierry
TITLE OF INVENTION: Isolated Tumor Rejection Antigen
TITLE OF INVENTION: Precursor Mage-2 Derived Peptides, and Uses Thereof
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSER: Felle & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/217,188A
FILING DATE: 24-MARCH-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5554724man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5340
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 838-3884
TELEFAX: (212) 838-9200
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
SEQUENCE CHARACTERISTICS:

```

```

LENGTH: 9 amino acid residues
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-217-188A-3

Query Match
Best Local Similarity 44.7%; Score 21; DB 1; Length 9;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

CY 4 MLDLH 9
Db 2 MVELVH 7

RESULT 11
US-08-687-226-3
Sequence 3, Application US/08687226
Patent No. 5686068
GENERAL INFORMATION:
APPLICANT: Melief, Cornelis J. M.; Visseren, M. W.;
APPLICANT: van der Burg, Goerd; van der Bruggen, Pierre;
APPLICANT: Boon-Falleur, Thierry
TITLE OF INVENTION: Isolated Peptides Derived From
TITLE OF INVENTION: Mage-2, Cytolytic T Cells Specific Of
TITLE OF INVENTION: Peptides And HLA-A2 Molecules, And Uses Thereof
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSER: Felle & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,226
FILING DATE: 25-JULY-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/217,188
FILING DATE: 24-MARCH-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5686068man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5447
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 838-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acid residues
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-687-226-3

Query Match
Best Local Similarity 44.7%; Score 21; DB 1; Length 9;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

CY 4 MLDLH 9
Db 2 MVELVH 7

RESULT 12
US-08-667-725B-3
Sequence 3, Application US/08667725B

```

Patent No. 6063900  
GENERAL INFORMATION:  
APPLICANT: Melief, Cornelis J. M.  
APPLICANT: Vissers, M. J. W.  
APPLICANT: Kast, W. M.  
APPLICANT: van der Bruggen, Pierre  
TITLE OF INVENTION: Isolated Tumor Rejection Antigen  
TITLE OF INVENTION: Precursor MAG-2 Derived Peptides, and Uses Thereof  
NUMBER OF SEQUENCES: 64  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski LLP  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/667,725B  
FILING DATE: 21 June 1996  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6063900man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5340.1 DIV (081585)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3000  
TELEFAX: (212) 752-5558  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acid residues  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-667-725B-3

Query Match 44.7%; Score 21; DB 3; Length 9;  
Best Local Similarity 50.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 MDLH 9  
DB 2 MWELVH 7

RESULT 13  
US-09-007-748-3  
Sequence 3, Application US/09007748  
Patent No. 6147187  
GENERAL INFORMATION:  
APPLICANT: Melief, Cornelis J. M.  
APPLICANT: Vissers, M. J. W.  
APPLICANT: Kast, W. M.  
APPLICANT: van der Bruggen, Pierre  
TITLE OF INVENTION: Isolated Tumor Rejection Antigen  
TITLE OF INVENTION: Precursor MAG-2 Derived Peptides, and Uses Thereof  
NUMBER OF SEQUENCES: 64  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski LLP  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/007,748  
FILING DATE: 15 January 1998  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6147187man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5340.2 DIV (081572)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3000  
TELEFAX: (212) 752-5558  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acid residues  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-007-748-3

Query Match 44.7%; Score 21; DB 3; Length 9;  
Best Local Similarity 50.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 MDLH 9  
DB 2 MWELVH 7

RESULT 14  
US-09-053-003-16  
Sequence 16, Application US/09053003  
Patent No. 6207391  
GENERAL INFORMATION:  
APPLICANT: Wu, Pengguang  
APPLICANT: McKiney, Judi  
TITLE OF INVENTION: High-Throughput Screening Assays for  
TITLE OF INVENTION: Modulators of STAT4 and STAT6 Activity  
NUMBER OF SEQUENCES: 64  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/053,003  
FILING DATE: 31-MAR-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Timothy L.  
REGISTRATION NUMBER: 35,367  
REFERENCE/DOCKET NUMBER: 018781-0008000US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 2

OTHER INFORMATION: /product= "OTHER"  
OTHER INFORMATION: /note= "Xaa = Tyr or phosphotyrosine"  
US-09-053-003-16

Query Match 44.7%; Score 21; DB 3; Length 9;  
Best Local Similarity 66.7%; Pred. No. 3.8e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 LDLLHY 10  
||:|  
Db 3 LDMPHV 8

RESULT 15  
US-09-183-931-31  
Sequence 31, Application US/09183931C

Patent No. 6210886  
GENERAL INFORMATION:  
APPLICANT: Van Baren, Nicolas  
APPLICANT: Brasseur, Francis  
APPLICANT: Boon-Falleur, Thierry  
TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING  
TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS  
FILE REFERENCE: LUD 5527.1-JEU/ES  
CURRENT APPLICATION NUMBER: US/09/183,931C  
CURRENT FILING DATE: 2000-02-28  
EARLIER APPLICATION NUMBER: US 09/018,422  
EARLIER FILING DATE: 1998 - 02 - 04  
NUMBER OF SEQ ID NOS: 44  
SEQ ID NO 31  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION:  
US-09-183-931-31

Query Match 44.7%; Score 21; DB 3; Length 9;  
Best Local Similarity 50.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 MDDLH 5  
||:|  
Db 2 MWELVH 7

Search completed: November 30, 2004, 08:15:58  
Job time : 28 secs